

UNIVERSIDAD AUTONOMA DE MADRID

ESCUELA POLITÉCNICA SUPERIOR



PROYECTO FIN DE CARRERA

# **Reconocimiento Facial Basado en Puntos Característicos de la Cara en entornos no controlados**

Ingeniería de Telecomunicación

Luis Blázquez Pérez

Enero 2013



# **Reconocimiento Facial Basado en Puntos Característicos de la Cara en entornos no controlados**

AUTOR: Luis Blázquez Pérez  
TUTOR: Pedro Tomé González

Área de Tratamiento de Voz y Señales (ATVS)  
Dpto. de Ingeniería Informática  
Escuela Politécnica Superior  
Universidad Autónoma de Madrid  
Enero de 2013



# Resumen

---

## *Resumen*

---

En este proyecto se estudia, implementa y evalúa un sistema automático de detección y corrección de puntos característicos faciales mal marcados, obtenidos mediante un sistema comercial automático. Como bases de datos para la experimentación se emplean varias bases de datos, emulando entornos controlados e incontrolados, de libre acceso a la comunidad científica. Se ha llevado a cabo un análisis antropométrico sobre el entorno controlado, probando su potencial; además de un análisis de cada uno de los rasgos faciales.

Como punto de partida se ha realizado un estudio de las técnicas biométricas desarrolladas en el estado del arte actual y una recopilación de aquellas con mayor interés. Debido a la existencia de ruido en las imágenes de ambas bases de datos, en este proyecto se ha desarrollado un método de preprocesado de imágenes. En él también se ha llevado a cabo el alineamiento, normalización (estándar ISO) y escalado de cada una de las imágenes.

Para la parte experimental se han llevado a cabo experimentos diferenciados de cada una de las etapas del sistema, de modo que se pueda evaluar el sistema desarrollado de forma detallada con sus puntos fuertes y débiles. Los experimentos se han centrado en observar el potencial del detector y del corrector de puntos faciales desarrollado en el proyecto, tanto en entornos controlados, como en entornos no controlados.

Además, para la realización de un análisis más exhaustivo de la cara, se han desarrollado dos segmentadores (uno basado en las proporciones faciales, y otro basado en el punto medio de cada rasgo facial), permitiendo así, la extracción de los rasgos faciales de manera individual, dando lugar a una mayor precisión en dicho análisis.

Por último, se presentan las conclusiones, las observadas de cada uno de los experimentos y las finales del proyecto, y se proponen líneas de trabajo futuras.

## *Palabras clave*

---

Sistema biométrico, reconocimiento biométrico facial, marcado de puntos característicos de la cara, detector de marcado incorrecto, corrector de puntos característicos, extractor de rasgos faciales.

### *Abstract*

---

This project is about study, implementation and evaluation of an automatic detection and correction system of misbranded facial landmarks, which are obtained from a commercial automatic system. Several different databases will be used for the experimentation, emulating both controlled and uncontrolled environments, which are free access for the scientific community. An anthropometric analysis about the controlled environment is performed, testing its potential; in addition to an analysis about every facial feature.

As starting point, a survey about the biometric techniques developed in the present state of the art and a compilation of the most interesting ones are made. Due to the existence of the noise in both databases images, an image preprocessing method is developed in this project. It also includes alignment, normalization (ISO standard) and escalade of each image.

In the experimental part different experiments are performed in every phase of the system, so it is possible to evaluate the system in detail, with its strengths and weaknesses. The experiments are focused on observing the potential of the detector and the facial points reviser, both in controlled and uncontrolled environments.

Furthermore, in order to perform a more exhaustive facial analysis, two segmentors have been developed (one of them based on facial proportions, and the another one based on the midpoint of each facial feature), thereby allowing the extraction of facial features individually, leading to a greater accuracy in this analysis.

Finally, conclusions about each experiment and about the entire project are presented, and future lines of work are suggested.

### *Key words*

---

Biometric system, facial biometric recognition, facial landmarks marking, inaccurate marking detector, landmarks reviser, facial features extractor.

# Agradecimientos

---

Quiero agradecer en primer lugar a mi ponente, el Dr. Julián Fierrez Aguilar, la oportunidad que me ha brindado de colaborar con el ATVS y su apoyo para realizar mi Proyecto Fin de Carrera.

Igualmente estoy muy agradecido al resto de miembros del ATVS, por haberme permitido disfrutar realizando este proyecto, y brindarme la oportunidad de iniciarme en el mundo de la investigación e impulsarme a continuar y ampliar aún más mis metas.

Especial agradecimiento por su dedicación a mi tutor Pedro Tomé González, que me ha guiado en la elaboración de este proyecto.

No quiero olvidar a mis amig@s, que durante todo el tiempo que ha durado mi carrera universitaria, la han hecho más amena y llevadera.

También quiero agradecer el incesante apoyo mostrado por mi familia a lo largo de mi vida. A mis hermanos agradecerles los momentos de diversión y distracción, que me ayudaban a desconectar cuando era necesario, del ámbito de trabajo y estudio. Papá, gracias por todo lo que has trabajado para darme lo mejor y por hacerme ver que todo en esta vida cuesta trabajo y que lo importante es luchar y valorar lo que se tiene. También te debo agradecer, a ti mamá, el haberme demostrado día tras día lo importante que soy en tu vida, dándolo todo sin pedir nada a cambio e inculcándome que no hay que rendirse en los malos momentos y que todo sucede por alguna razón. Por todo ello, por hacer que esta vida fuera mucho más fácil para mí, y sobre todo, porque yo no sería lo que soy sin vosotros, GRACIAS FAMILIA.

Y en último lugar y no por ello menos importante, quiero dedicar este proyecto a Patricia Domenech, que es capaz de iluminarme en cualquier día de oscuridad gracias a su energía, sonrisa y vitalidad.

Gracias a todos por hacer de estos años una época memorable.

*Luis Blázquez Pérez  
Enero 2013*





A mis padres, por ayudarme cada día a  
cruzar con firmeza el camino de la superación.

El genio comienza las grandes obras,  
pero solo el trabajo las acaba.

Petrus Jacobus Joubert

# Índice de contenidos

<b>Resumen</b> . . . . .	<b>III</b>
<b>Palabras claves</b> . . . . .	<b>III</b>
<b>Abstract</b> . . . . .	<b>IV</b>
<b>Key words</b> . . . . .	<b>IV</b>
<b>Agradecimientos</b> . . . . .	<b>V</b>
<b>Índice de contenidos</b> . . . . .	<b>VII</b>
<b>Índice de figuras</b> . . . . .	<b>IX</b>
<b>Índice de tablas</b> . . . . .	<b>XI</b>
<b>Glosario de términos</b> . . . . .	<b>XII</b>
<b>1. Introducción</b> . . . . .	<b>1</b>
1.1. Motivación. . . . .	1
1.2. Objetivos del PFC . . . . .	3
1.3. Metodología y plan de trabajo. . . . .	4
1.4. Organización de la memoria . . . . .	5
<b>2. Estado del arte en localización de puntos característicos de la cara</b> . . . . .	<b>6</b>
2.1. Introducción. . . . .	6
2.2. Métodos de localización de puntos característicos en la cara. . . . .	6
2.2.1. Correspondencia entre agrupaciones de grafos elásticos (Elastic Bunch Graph Matching, EBGM) . . . . .	7
2.2.2. Modelo de apariencia activa (Active Appearance Model, AAM). . . . .	8
2.2.3. Modelo de forma activa (Active Shape Model, ASM) . . . . .	10
<b>3. Bases de datos</b> . . . . .	<b>13</b>
3.1. Introducción. . . . .	13
3.2. BioID . . . . .	13
3.3. Forensic ATVS . . . . .	14
3.4. MORPH. . . . .	17
3.5. SCFace . . . . .	19
<b>4. Sistema, diseño y desarrollo</b> . . . . .	<b>23</b>
4.1. Introducción a los sistemas biométricos . . . . .	23
4.1.1. Modo identificación . . . . .	24
4.1.2. Modo verificación . . . . .	25
4.1.3. Otras consideraciones . . . . .	26
4.2. Sistema automático comercial. . . . .	26
4.3. Descripción del sistema desarrollado . . . . .	27
4.4. Pre-procesado y normalización . . . . .	28
4.5. Detector de marcado incorrecto. . . . .	32
4.5.1. Fase 1. No landmarks . . . . .	34
4.5.2. Fase 2. Marcado incorrecto de ojos. . . . .	34
4.5.3. Fase 3. Distancias. . . . .	37
4.5.4. Fase 4. Ángulos . . . . .	37
4.5.5. Fase 5. Simetrías . . . . .	37
4.5.6. Decisor . . . . .	41
4.6. Corrector de puntos característicos . . . . .	43
4.6.1. Extractor de rasgos faciales por proporciones. . . . .	44
4.6.2. Extractor de rasgos faciales por puntos . . . . .	46
4.6.3. Máscaras de ruido . . . . .	48
4.6.4. Realce, ecualización y umbralización . . . . .	49
4.6.5. Barrido, localización y marcado de puntos. . . . .	51

<b>5. Experimentos y resultados . . . . .</b>	<b>56</b>
5.1. Evaluación del rendimiento de sistemas biométricos . . . . .	56
5.2. Análisis de características / puntos característicos. . . . .	57
5.3. Experimento 0: Marcado automático de puntos característicos en entornos controlados . . . . .	62
5.4. Experimento 1: Importancia de la segmentación . . . . .	66
5.5. Experimento 2: Reconocimiento antropométrico . . . . .	67
5.5.1. Protocolo experimental . . . . .	67
5.5.2. Resultados y conclusiones. . . . .	68
5.6. Experimento 3: Análisis de poder discriminante de cada rasgo facial para ambos tipos de extractores. . . . .	74
5.6.1. Protocolo experimental . . . . .	74
5.6.2. Resultados . . . . .	75
5.6.3. Discusión . . . . .	79
<b>6. Conclusiones y trabajo futuro . . . . .</b>	<b>80</b>
6.1. Conclusiones. . . . .	80
6.2. Trabajo futuro. . . . .	81
<b>7. Referencias bibliográficas. . . . .</b>	<b>83</b>
<b>A. Presupuesto . . . . .</b>	<b>85</b>
<b>B. Pliego de condiciones . . . . .</b>	<b>86</b>
<b>C. Publicaciones. . . . .</b>	<b>90</b>

# Índice de figuras

Figura 1. Facial landmarks . . . . .	3
Figura 2. Correspondencia entre agrupaciones de grafos elásticos. . . . .	7
Figura 3. Ejemplo de malla en AAM . . . . .	9
Figura 4. Búsqueda de una cara utilizando el Modelo de Forma Activa . . . . .	11
Figura 5. Cuatro usuarios, a modo de ejemplo, de la base de datos BioID. . . . .	13
Figura 6. Marcado de puntos BioID . . . . .	14
Figura 7. Distancias de captura en la base de datos ATVS . . . . .	14
Figura 8. Imágenes ejemplo de la base de datos ATVS a diferentes distancias . . . . .	15
Figura 9. Conjunto de imágenes para un único usuario . . . . .	15
Figura 10. Puntos faciales Forensic ATVS . . . . .	17
Figura 11. Cuatro usuarios, a modo de ejemplo, de la base de datos MORPH. . . . .	18
Figura 12. Imagen progresión edad hombre blanco (izq.) y mujer Afro-americana (dcha.). .	19
Figura 13. Posición cámaras comerciales de calidad variable . . . . .	20
Figura 14. Imagen frontal (izquierda) e imagen camara1 (derecha) de dos usuarios de SCface . . . . .	20
Figura 15. Cámaras de vigilancia. . . . .	20
Figura 16. Distancias de captura de las imágenes de vigilancia . . . . .	21
Figura 17. Ejemplo del conjunto de imágenes por usuario . . . . .	22
Figura 18. Distribución de la edad de los sujetos de la base de datos . . . . .	22
Figura 19. Módulo de inscripción . . . . .	23
Figura 20. Diagrama de bloques de un sistema biométrico de reconocimiento. . . . .	24
Figura 21. Esquema modo identificación . . . . .	24
Figura 22. Esquema modo verificación . . . . .	25
Figura 23. Puntos faciales obtenidos a través del FaceSDK . . . . .	26
Figura 24. Esquema general para la detección y corrección de puntos característicos de la cara. . . . .	27
Figura 25. Etapa de pre-procesado . . . . .	28
Figura 26. Marcado de las pupilas original (izquierda) y después del pre-procesado (derecha). . . . .	29
Figura 27. Formato geométrico (izquierda) y un ejemplo con W=240 píxeles (derecha). . .	30
Figura 28. Imagen original (izquierda) e imagen normalizada y rotada (derecha) . . . . .	31
Figura 29. Transformación Afín y recorte . . . . .	32
Figura 30. Bloque detector . . . . .	32
Figura 31. Detector de puntos característicos . . . . .	33
Figura 32. Elementos estructurantes planos . . . . .	34
Figura 33. Representación fase 2 . . . . .	35
Figura 34. Ejemplo imágenes con mal marcado de ojos (ojos en parte lisa de la cara) . . .	35
Figura 35. Restricciones de la etapa 2, fase 2. . . . .	36
Figura 36. Ejemplo imagen marcada de forma correcta . . . . .	36
Figura 37. Distancias calculadas . . . . .	37
Figura 38. Ángulos. . . . .	37
Figura 39. Distancias a la vertical (izquierda) y distancias a la horizontal (derecha). . . . .	39
Figura 40. Histogramas resultado para ángulos $\alpha$ . . . . .	40
Figura 41. Ejemplo de imágenes que fallan en cada una de las fases del detector, para la base de datos MORPH . . . . .	41
Figura 42. Esquema decisor. . . . .	41
Figura 43. Imágenes mal marcadas de MORPH. . . . .	42
Figura 44. Bloque corrector de puntos faciales . . . . .	43

Figura 45. Funcionamiento corrector puntos característicos . . . . .	43
Figura 46. Proporciones faciales utilizadas . . . . .	45
Figura 47. Resultado del segmentador basado en proporciones faciales . . . . .	45
Figura 48. Resultado segmentador por puntos (ATVS). . . . .	47
Figura 49. Comparación resultado de ambos segmentadores. . . . .	47
Figura 50. Máscaras elípticas. . . . .	48
Figura 51. Máscaras triangulares . . . . .	48
Figura 52. Máscaras dos pendientes. . . . .	49
Figura 53. Realce . . . . .	49
Figura 54. Ecualización . . . . .	50
Figura 55. Umbralización . . . . .	50
Figura 56. Barrido vertical . . . . .	51
Figura 57. Resultado algoritmo barrido vertical sobre ojo. . . . .	51
Figura 58. Resultado algoritmo barrido vertical + método especial para la boca . . . . .	53
Figura 59. Comparativa puntos característicos de la boca. . . . .	53
Figura 60. Comparación punto característico de la barbilla . . . . .	54
Figura 61. Puntos característicos sobre un usuario de la base de datos MORPH . . . . .	54
Figura 62. Base de datos MORPH corregida. . . . .	55
Figura 63. Equal Error Rate. . . . .	56
Figura 64. Puntos característicos utilizados por la DGGC . . . . .	58
Figura 65. Mediciones a tomar. . . . .	59
Figura 66. Puntos característicos de la cara de interés en el proyecto . . . . .	60
Figura 67. Resultados de marcado entre los puntos proporcionados y el sistema desarrollado en la base de datos BioID. . . . .	63
Figura 68. Error medio, de cada rasgo facial, por usuario . . . . .	64
Figura 69. Error medio de cada usuario . . . . .	65
Figura 70. Usuarios con rasgos faciales ocluidos . . . . .	65
Figura 71. Vectores de características. . . . .	68
Figura 72. Distancias más discriminantes de la cara . . . . .	70
Figura 73. Error de segmentación ATVS. . . . .	71
Figura 74. Vectores de características rasgos faciales . . . . .	72
Figura 75. Distancias más discriminantes de cada rasgo. . . . .	73
Figura 76. Usuario de la base de datos ATVS . . . . .	75
Figura 77. Rasgos faciales extraídos. . . . .	76
Figura 78. Error medio por rasgos y distancia. . . . .	76
Figura 79. EER rasgos faciales mediante texturas para ambos segmentadores ordenados de mejor a peor . . . . .	78

# Índice de tablas

---

Tabla 1. Descripción bases de datos. . . . .	13
Tabla 2. Número de Imágenes por género y ascendencia (n=55134) . . . . .	18
Tabla 3. Número de imágenes por usuario . . . . .	18
Tabla 4. Número de Imágenes por género y edad (n=55134). . . . .	18
Tabla 5. Detalles de los metadatos . . . . .	19
Tabla 6. Especificaciones de las cámaras de vigilancia. . . . .	21
Tabla 7. Especificaciones cámara utilizada para imágenes frontales . . . . .	21
Tabla 8. Características geométricas de las imágenes . . . . .	30
Tabla 9. Tabla de proporciones (el color de la letra hace referencia al color de las líneas). . . . .	44
Tabla 10. Offsets para las imágenes normalizadas, en píxeles . . . . .	46
Tabla 11. Grado de semejanza. . . . .	57
Tabla 12. Puntos característicos utilizados por la DGGC. . . . .	58
Tabla 13. Puntos de interés faciales. . . . .	60
Tabla 14. Longitud vectores de características. . . . .	61
Tabla 15. Error sistema manual vs sistema automático . . . . .	64
Tabla 16. Error medio de cada rasgo . . . . .	66
Tabla 17. Número de imágenes erróneas. . . . .	66
Tabla 18. Error de segmentación . . . . .	67
Tabla 19. Comparación EER . . . . .	69
Tabla 20. Distancias por rasgos . . . . .	72
Tabla 21. EERs clasificados por segmentador y bases de datos. . . . .	79

# Glosario de términos

---

AAM: Active Appearance Model (Modelo de apariencia activa)

ASM: Active Shape Model (Modelo de forma activa)

CCTV: Circuito Cerrado de Televisión

DET: Detection Error Trade-off (compensación por error de detección)

DGGC: Dirección General de la Guardia Civil

DNI: Documento Nacional de Identidad

DPC: Data Protection Commission (Comisión de Protección de Datos)

EBGM: Elastic Bunch Graph Matching (Correspondencia entre agrupaciones de grafos elásticos)

EE: Elemento Estructurante

EER: Equal Error Rate (tasa de error igual)

FA: Falsa Aceptación

FAR: False Acceptance Rate (tasa de falsa aceptación)

FR: Falso Rechazo.

FRR: False Rejection Rate (tasa de falso rechazo)

IPD: Interpupillary Pixel Distance (Distancia interpupilar)

IR: Infrarrojos

LDA: Linear discriminant analysis (Análisis discriminante lineal)

PFC: Proyecto de Fin de Carrera

PCA: Principal Components Analysis (Análisis de Componentes Principales)

SFFS: Sequential Forward Floating Selection (Selección secuencial hacia delante)

SVM: Support Vector Machine (Máquina de Vectores de Soporte)





# 1. Introducción

---

## 1.1 Motivación

---

La biometría se encarga del reconocimiento automático de personas mediante sus rasgos físicos (cara, retina, iris, voz, huellas dactilares, etc.) o rasgos de conducta (forma de andar, de escribir, etc.).

En principio, cualquier característica física o determinados tipos de comportamiento pueden ser utilizados como rasgos biométricos, siempre y cuando posean las siguientes propiedades:

- a. Universalidad: toda persona debe poseer la característica biométrica utilizada.
- b. Capacidad de distinción: las características biométricas de dicho rasgo, debe permitir discriminar a dos personas distintas.
- c. Constancia: debe permanecer relativamente invariante a lo largo del tiempo.
- d. Capacidad de cuantificación: debe poder ser medida.

En la actualidad existen tres métodos de reconocimiento facial:

- De rasgos locales: reconocen los ojos, la nariz, la boca, miden las distancias y los ángulos de la cara.
- De rasgos globales: aportan información de toda la cara.
- Mixtos: combinación de los anteriores.

El sistema de rasgos globales, tiene un margen de error más pequeño, aunque maneja un menor nivel de detalle.

Una de las principales ventajas del reconocimiento facial, es que se trata de un método no intrusivo, es decir, los datos pueden ser adquiridos incluso sin que el sujeto se percate de ello. Además, el aspecto facial es el método más utilizado de manera natural por los seres humanos para reconocerse unos a otros.

Sin embargo, a la hora de identificar a una persona a partir de su aspecto facial, existen una serie de dificultades centradas, sobretudo, en el concepto de variabilidad. Se hace muy difícil el reconocimiento facial cuando hay la variabilidad entre individuos es muy pequeña (por ejemplo, es el caso de familiares o, especialmente, es el caso de los gemelos), o cuando la variabilidad entre distintas imágenes de un mismo individuo es muy amplia; esto puede ser debido a que dichas imágenes hayan sido adquiridas en diferentes condiciones de posición o iluminación. Otra de los grandes obstáculos en el reconocimiento facial, es la base de datos, esto es, para poder evaluar los métodos de reconocimiento facial, la base de datos debe ser adquirida en condiciones equivalentes para todos los individuos y en número de personas suficientemente representativo, pero teniendo en cuenta que un aumento en el número de datos, aumenta el gasto computacional.

En la actualidad, existen diversas aplicaciones cuyo funcionamiento está basado en el reconocimiento facial, dividiéndose en tres grandes grupos.

El grupo más conocido por la gente, es el **comercial**, es decir, aplicaciones para la introducción de un usuario en redes de ordenadores, seguridad electrónica, acceso a internet, cajeros automáticos, controles de acceso, teléfonos móviles, etc. Entre las aplicaciones que tienen cabida en este primer grupo, son remarcables los ejemplos de Google y Facebook.

La última innovación de Google, ha sido un nuevo sistema facial que permite al usuario el acceso a la información de su Smartphone o tableta. Dicho reconocimiento facial permite crear a la vez distintos perfiles en un mismo aparato. Solo hace falta que registre sus datos biométricos para que sus dispositivos los reconozca, después de configurar el acceso a sus datos y documentos personales para que solo el usuario tenga acceso a ellos. La localización de los rasgos biométricos de cada persona se basaría en un análisis de la forma y el tamaño de la nariz, la mandíbula y los pómulos. Google ha patentado este desbloqueo de dispositivos con un simple reconocimiento facial<sup>1</sup>.

Por otra parte, Facebook ha creado una herramienta de reconocimiento facial, que permite reconocer al usuario mediante una foto. Sin embargo, en este caso, han tenido que desactivar su tecnología de reconocimiento facial y borrar los datos incluidos en los perfiles de sus usuarios, debido a la Comisión de Protección de Datos (DPC), ya que exigen una mayor transparencia para los usuarios sobre la manera en que se tratan sus datos.

El segundo gran grupo al que se hace referencia, es el **gubernamental**, cuyas aplicaciones son las vinculadas a documentos identificativos (DNI, pasaporte, permiso de conducir), Seguridad Social, control de fronteras, control de aeropuertos, etc. La aplicación más conocida y moderna, perteneciente a este segundo grupo es el ABC system, utilizado para el chequeo de pasaportes (se pone un pasaporte y se comprueba que la foto coincide con la del mismo, registrada en la base de datos). Este sistema se ha implementado en el Aeropuerto de Barajas, como un nuevo método de reconocimiento facial. A partir de la imagen de la cara, se determinan unos cuantos números que representen únicamente a esa persona y se compara con los de una base de datos para conocer quién es ese individuo. Para obtener la información adecuada, lo mejor es que la imagen se capte lo más frontal posible. El método de reconocimiento facial implementado es el de rasgos globales, debido a que se da más prioridad a la falta de errores que al nivel de detalle. En [1] y [2] se explican las directrices técnicas y operacionales, respectivamente.

En el tercer grupo, el **forense**, tienen lugar investigaciones criminales, identificación de cadáveres, terrorismo, identificación de personas desaparecidas, etc.

---

<sup>1</sup> <http://eldia.es/2012-09-07/sociedad/0-Google-desarrolla-desbloqueo-facial-varios-usuarios-mismo-telefono.htm>

En un sistema biométrico real, además del porcentaje de aciertos o fallos, deben existir otras cuestiones que han de ser consideradas:

- a. Funcionamiento: el porcentaje de reconocimiento debe superar un determinado umbral y el tiempo de ejecución debe ser el menor posible.
- b. Aceptabilidad: debe ser aceptado por las personas como un método de reconocimiento que puedan integrar en su vida diaria.
- c. Robustez: debe ser estable frente a ataques fraudulentos.

### *1.2 Objetivos del PFC*

---

Los objetivos del presente proyecto son la segmentación de la cara y el análisis exhaustivo de cada uno de los rasgos faciales extraídos.

Con los bloques desarrollados, para la cumplimentación de dichos objetivos, detector, corrector y extractor, se pretende realizar un sistema que compruebe la validez de sistemas comerciales basados en el marcado de puntos característicos de la cara (en el caso de este proyecto, el sistema comercial que se pondrá a prueba es el llamado FaceSDK), y que en caso de fallo, los corrija de forma fiable.

La utilidad de estos bloques es muy alta, ya que la correcta estimación de los puntos faciales de interés (Figura 1) tiene numerosas ventajas, como la extracción de rasgos faciales de manera más precisa y exacta o la realización de un análisis de morfología (permitiendo el estudio de texturas de los rasgos).

- 1 = right mouth corner
- 2 = left mouth corner
- 3 = outer end of right eye brow
- 4 = inner end of right eye brow
- 5 = inner end of left eye brow
- 6 = outer end of left eye brow
- 7 = outer corner of right eye
- 8 = inner corner of right eye
- 9 = inner corner of left eye
- 10 = outer corner of left eye
- 11 = right nostril
- 12 = left nostril
- 13 = tip of chin



Figura 1. Facial landmarks.

Finalmente, se estudiarán las principales distancias entre puntos característicos, para el reconocimiento, mediante los últimos dos bloques desarrollados, utilizados para clasificar y comparar los vectores de características de las imágenes, además de analizar el poder discriminante de cada rasgo, gracias a los extractores desarrollados.

### 1.3 Metodología y plan de trabajo

---

La metodología de trabajo se divide en los siguientes pasos:

#### **Estudio de la literatura**

- El primer paso será un estudio de los sistemas de reconocimiento facial, en particular, los utilizados para el marcado de puntos característicos de la cara.
- Aquí se revisa el estado del arte en dichos sistemas [3,4].
- Para completar la fase de estudio, las técnicas básicas de procesamiento de imagen digital serán estudiadas a través de la comprensión del libro [5] y la realización de algunos de sus ejemplos.

**Desarrollo de un sistema de detección y marcado/corrección de puntos característicos faciales** [6,7,8,9]. Este sistema se aplica tanto en entornos controlados como no controlados. El desarrollo incluye los siguientes pasos:

- Preprocesamiento de las imágenes.
- Detección de los rasgos biométricos (ojos, nariz, boca...) mal marcados.
- Extracción de dichos rasgos biométricos [10]
- Corrección de los puntos característicos.

**Desarrollo de un sistema de identificación biométrica.** Este sistema utilizará las características extraídas de las imágenes generadas en la etapa anterior. El desarrollo incluye los siguientes pasos.

- Extracción de características (distancias euclídeas entre los puntos característicos, ya corregidos).
- Medida de similitud
- Toma de decisiones.

**Experimentos para la optimización del sistema.** Estos experimentos se centran en la selección de características. Se llevan a cabo para obtener un conjunto de características distintivas y representante de cada usuario para aumentar el rendimiento del sistema.

**Evaluación del sistema mediante la realización de diferentes experimentos.**

**Evaluación de los resultados y las conclusiones extraídos.**

**Documentación de los trabajos realizados.**

- Descripción de los pasos seguidos.
- Análisis y evaluación de los resultados.
- Posible trabajo futuro.

### ***1.4 Organización de la memoria***

---

El presente documento está estructurado en 6 capítulos:

**Capítulo 1: Introducción.** Éste capítulo se presenta el tema de este proyecto, las principales razones que nos han animado a desarrollar este trabajo, los objetivos a conseguir, la metodología seguida y la organización de la memoria.

**Capítulo 2: Estado del arte en localización de puntos característicos de la cara.** En este capítulo se examinan algunas de las actuales tecnologías de marcado de puntos característicos de la cara: 1) EBGM, 2) AAM, 3) ASM, así como sus aplicaciones biométricas.

**Capítulo 3: Bases de datos.** En este capítulo se describen cuatro bases de datos de imágenes: 1) BioID [11], que es una base de datos generada en entornos controlados, 2) Forensic ATVS, que es la base de datos generada durante el desarrollo de este proyecto en condiciones controladas, 3) MORPH [12], que es una base de datos obtenida en entornos semi-controlados y 4) Scface [13], que es una base de datos generada en entornos no controlados y con una calidad de imagen mala.

**Capítulo 4: Sistema, diseño y desarrollo.** En el cuarto capítulo, después de una breve introducción a los sistemas biométricos, se detallan las diferentes etapas del diseño y desarrollo del sistema biométrico implementado: desde el pre-procesamiento de las imágenes a la decisión de la identidad.

**Capítulo 5: Experimentos y resultados.** En este capítulo, en primer lugar, se presenta un resumen de la evaluación del desempeño de los sistemas biométricos. Entonces, se describen los protocolos experimentales y los clasificadores empleados en los experimentos. Finalmente, se presentan y analizan los resultados obtenidos en cada experimento.

**Capítulo 6: Conclusiones y trabajo futuro.** Las conclusiones, juntos con el posible trabajo futuro, son expuestas en la parte final del proyecto.

Por último, las secciones siguientes se encuentran al final de la disertación: la bibliografía, el presupuesto del proyecto, el pliego de condiciones y el artículo publicado.

## 2. Estado del arte en localización de puntos característicos de la cara

---

### 2.1 Introducción

---

El reconocimiento facial automatizado es un concepto relativamente nuevo, pues se introdujo en los años 60. Fue entonces cuando se desarrolló el primer sistema semiautomático para reconocimiento facial, el cual requería la imagen de una persona para localizar los rasgos (como ojos, nariz y boca) en las fotografías antes de que este calculara distancias a puntos de referencia en común, que posteriormente eran comparados con datos de referencia [14].

En los años 70 Goldstein, Harmon & Lesk [15], utilizaron 21 marcadores subjetivos específicos tales como el color del cabello y el grosor de labios para automatizar el reconocimiento facial. El problema con estas soluciones previas era que se seguían requiriendo un proceso manual. En 1988 Kirby & Sirobich aplicaron análisis de componentes principales (PCA), una técnica estándar del álgebra lineal, al problema del reconocimiento facial, para aumentar la exactitud de los resultados. Esto fue considerado un avance muy importante al mostrar que eran requeridos menos de 100 valores para codificar acertadamente la imagen de una cara convenientemente alineada y normalizada [16].

En 1991 Turk & Pentland, utilizando las técnicas de Eigenfaces, como se denominó al método de Kirby & Sirobich, demostraron que el error residual podía ser utilizado para detectar caras en las imágenes [17], un descubrimiento que permitió desarrollar sistemas automatizados fiables de reconocimiento facial en tiempo real. Si bien la aproximación era un tanto forzada por factores ambientales, creó sin embargo un interés significativo en posteriores desarrollos de éstos sistemas.

### 2.2 Métodos

---

Hay tres enfoques predominantes en el problema de reconocimiento facial: i) geométrico: basado en los rasgos, ii) fotométrico: basado en lo visual, y iii) la combinación de los dos enfoques anteriores.

Conforme a que el interés en reconocimiento facial continuó, fueron desarrollados muchos algoritmos diferentes; se deben destacar tres de ellos, los cuales han sido bien estudiados en la literatura del reconocimiento facial:

- Correspondencia entre agrupaciones de grafos elásticos (Elastic Bunch Graph Matching, EBGM),
- Modelo de apariencia activa (Active Appearance Model, AAM),
- Modelo de forma activa (Active Shape Model, ASM).

Los sistemas de reconocimiento facial modernos se basan en estos algoritmos.

### 2.2.1 Correspondencia entre agrupaciones de grafos elásticos (Elastic Bunch Graph Matching, EBGM)

El algoritmo EBGM [18] fue diseñado para realizar tareas de reconocimiento facial, utilizando solamente algunos puntos de interés y no la cara en su totalidad. La técnica de EBGM se desarrolla básicamente en dos etapas: la primera consiste en ajustar un grafo de puntos principales a la cara del individuo, utilizando para ello un modelo estadístico de dicho grafo; la segunda etapa extrae características locales en dichos puntos y halla la distancia entre el grafo obtenido y sus descriptores al grafo almacenado de la persona a identificar. Dependiendo de la distancia encontrada, se ratifica o no la identidad del individuo.

El primer paso, antes de la utilización de este algoritmo, es la normalización de las imágenes. En este proceso se reduce el tamaño de las imágenes y se realizan transformaciones geométricas que ubican las coordenadas de los ojos del individuo en posiciones predeterminadas.

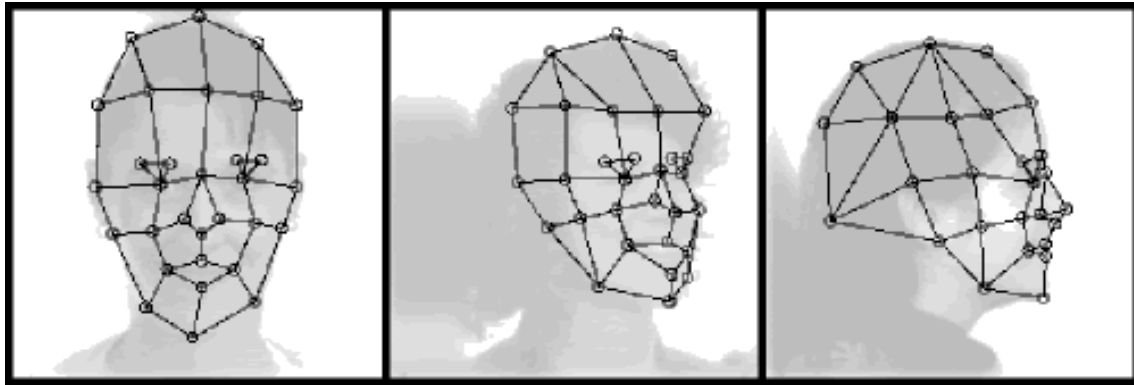


Figura 2. Correspondencia entre agrupaciones de grafos elásticos

La implementación del algoritmo EBGM (Figura 2) requiere el uso de las wavelets Gabor (filtros espaciales paso-banda, que permiten alcanzar la resolución conjunta de información máxima en los espacios bidimensionales espacial y frecuencial), ya que estas brindan una descripción de la información en frecuencia, en una región específica de la imagen. Para ello es necesario realizar la operación de convolución de la región con una variedad de wavelets; estas wavelets son almacenadas como máscaras en arreglos bidimensionales. Las máscaras se pueden definir mediante la ecuación 1, donde las wavelets se separan en parte par y parte impar.

$$\text{Par. } Ond_p(x, y) = e^{-\frac{(x-x_0)^2 + \gamma^2(y-y_0)^2}{2\sigma^2}} * \cos\left(\frac{2 * \pi * x_r}{\lambda}\right) \quad (1)$$

$$\text{Impar. } Ond_i(x, y) = e^{-\frac{(x-x_0)^2 + \gamma^2(y-y_0)^2}{2\sigma^2}} * \sin\left(\frac{2 * \pi * x_r}{\lambda}\right)$$

Con:  $x_r = (x-x_0) * \cos(\theta) + (y-y_0) * \sin(\theta)$ ; donde  $\lambda$  define la longitud de onda,  $\sigma$  define el radio de la gaussiana,  $\gamma$  define la relación de aspecto de la gaussiana,  $\theta$  representa la orientación de la portadora y  $(x_0, y_0)$  determinan la ubicación del valor pico de la función gaussiana.

A continuación se describe la parte del algoritmo EBGM que se refiere a la localización de puntos característicos.

1. Se define una estructura de grafo sobre la cara, cuyos nodos son puntos de interés, que se puedan localizar fácilmente y que posean la misma estructura en todos los rostros.
2. Cada nodo del grafo es caracterizado utilizando un banco de filtros de Gabor de diferentes frecuencias y orientaciones. En cada nodo, se calcula la respuesta de todos los filtros, denominado Jet. Por consiguiente, cada nodo queda etiquetado con sus coordenadas y su Jet asociado.
3. Para una imagen nueva se busca en la base de datos el conjunto de Jets que sean más similares. Para realizar esto, se centra el grafo en los ojos de la nueva imagen, se calculan los jets de estos puntos y se comparan con los jets de los ojos de la base de datos. Para la comparación de los jets se utiliza una función de similitud de fase, ésta es similar a una correlación, y está definida como se describe en la ecuación 2.

$$S(J, J') = \frac{\sum_{j=1}^n a_j a'_j \cos(\phi_j - \phi'_j)}{\sqrt{\sum_{j=1}^n a_j^2 \sum_{j=1}^n a'^2_j}} \quad (2),$$

donde:  $a'$  es la magnitud del jet de la base de datos,  $a$  la magnitud del Jet de la imagen nueva,  $\phi'_j$  la fase del Jet de la base de datos,  $\phi_j$  la fase del Jet de la imagen nueva y  $n$  el número de coeficientes del Jet.

Este algoritmo tiene en cuenta que las imágenes faciales reales tienen muchas características no lineales que no son tratadas en los métodos lineales de análisis (LDA, PCA), tales como variaciones en la iluminación (iluminación de exteriores vs interior fluorescente), postura (frontal vs inclinada) y expresión (sonrisa vs ceño fruncido). Sin embargo, la dificultad de este método es el requerimiento de la precisa localización del punto de referencia.

### 2.2.2 Modelo de apariencia activa (Active Appearance Model, AAM)

Los Modelos de Apariencia Activa (AAM) [19] permiten reproducir de forma sintética imágenes de superficies que incluyen deformaciones no rígidas y cambios de apariencia. Están basados en la obtención, mediante una fase de entrenamiento, de un modelo estadístico de la forma y la apariencia del objeto de interés.

En un AAM la forma es descrita mediante un conjunto de  $N$  puntos característicos (los llamados *Parámetros de forma*), que determinan una malla similar a la representada en la Figura 3 y que es expresada por el siguiente vector:

$$s = (u_1, v_1, u_2, v_2, \dots, u_N, v_N) \quad (1),$$

donde  $u_i, v_i$  son las coordenadas del vértice  $i$ .





Figura 3. Ejemplo de malla en AAM

Mediante el análisis de componentes principales (PCA) sobre las mallas de entrenamiento se obtiene una malla  $s_0$  y un subespacio  $B_s=[s_1,...,s_N]$  formado por  $n$  componentes principales, con una dimensionalidad menor que la del conjunto de entrenamiento.

Cualquier instancia de la forma del modelo se obtiene a partir de una combinación lineal de los vectores de la base de forma,  $B_s$  mediante la siguiente expresión:

$$s(p) = s_0 + \sum_{i=1}^n p_i s_i \quad (2)$$

La inclusión de los parámetros de forma en el modelo se realiza mediante una transformación afín definida a trozos denominada función warp  $W(x;p)$ . Esta función se encarga de transformar los puntos interiores de una malla concreta (normalmente se elige  $s_0$ ), en donde se define la apariencia, a cualquier malla  $s(p)$  generada a partir de (2). Es decir:

$$x' = W(x;p) \quad (3),$$

Donde  $x$  son puntos en el interior de  $s_0$  y  $x'$  está definido en el interior de  $s(p)$ .

La apariencia se describe a partir del mapa de bits definido en el interior de los diversos triángulos que forman los puntos de la malla  $s_0$ , son los llamados *Parámetros de apariencia*.

Mediante (3) se transforman las imágenes de entrenamiento, con el fin de normalizarlas en forma. De la misma manera que con los parámetros de forma, mediante PCA se obtienen tanto la apariencia media  $A_0$ , como la base de un subespacio  $B_A=[A_1(x), A_2(x),...,A_m(x)]$ , de dimensión menor al conjunto de entrenamiento, que está formada por las  $m$  componentes principales del entrenamiento.

A partir de estos elementos, se obtiene un modelo de apariencia lineal, que es capaz de generar una instancia de apariencia a partir de una combinación lineal de la media y las componentes de la base, ponderadas por un conjunto de parámetros  $\lambda=(\lambda_1,\lambda_2,...,\lambda_m)$ :

$$A(x;\lambda) = A_0(x) + \sum_{i=1}^m \lambda_i A_i(x) \quad (4)$$

Una vez obtenido el modelo de apariencia lineal, se procede con el *ajuste del modelo* que trata de, a partir de una imagen de entrada  $I(x)$ , encontrar el conjunto de parámetros  $p$  y  $\lambda$  que minimicen el error cuadrático entre la instancia del modelo generado a partir de esos parámetros y la imagen de entrada:

$$\sum_x [A_0(x) + \sum_{i=1}^m \lambda_i A_i(x) - I(W(x; p))]^2 \quad (5)$$

Existen diversos métodos para minimizar (5), entre los que destacan por su precisión el algoritmo de Lucas-Kanade, el cual, en líneas generales, es un método de minimización iterativo basado en el algoritmo de Gauss-Newton.

El modelo AAM es un buen método estadístico para ajustes de plantillas, el cual usa toda la información de la cara, es decir, no solo los puntos de los bordes, sino también los interiores. El gran problema de este modelo, es que son muy sensibles al proceso de inicialización, esto es, cuando la inicialización está alejada del objeto destino, ellos pueden converger hacia mínimos locales.

### 2.2.3 Modelo de forma activa (Active Shape Model, ASM)

El modelo estructural que define ASM [20] está compuesto por tres elementos: i) un modelo de contorno, ii) un modelo de apariencia y iii) un algoritmo de búsqueda para ajustar el modelo mediante la minimización de una función de coste.

#### *i. Modelo de contorno*

El objetivo del modelo de contorno es asegurar que la segmentación produce contornos válidos para la región facial. El modelo de contorno viene dado por los componentes principales de los vectores definidos por las marcas de referencia. Una región facial queda definida mediante  $n$  marcas  $(x_1, y_1), \dots, (x_n, y_n)$  que se pueden agrupar para formar vectores de contorno de la forma:

$$x = (x_1, \dots, x_n, y_1, \dots, y_n)^T \quad (1)$$

Así, un contorno se puede aproximar mediante la siguiente expresión:

$$x \approx x_m + \phi * b \quad (2)$$

En la expresión anterior,  $x_m$  define el contorno medio,  $\phi$  es la matriz de  $t$  autovectores obtenida como solución al problema de autovectores sobre la matriz de covarianzas de los contornos y  $b$  es un vector de  $t$  elementos que contiene los parámetros del modelo tal que:

$$b = \phi^T * (x - x_m) \quad (3)$$

Cuando se ajusta el modelo, los valores de  $b$  se limitan de la siguiente forma:

$$|b_i| < m * \sqrt{\lambda_i} \quad (4)$$

De manera general, los valores de  $m$  oscilan entre dos y tres. Por su parte, el número de autovalores que se emplean, se eligen de tal forma que el modelo ajuste una cierta proporción  $f_v$  de la varianza de los contornos de entrenamiento.

## ii. Modelo de Apariencia

El modelo de apariencia se limita al borde de la región facial sobre la imagen en escala de grises y es la derivada normalizada en la dirección perpendicular al contorno de la misma. Su objetivo es asegurar que la segmentación localiza el rostro en una posición donde la estructura de la imagen a los lados (o dentro) del contorno es similar a la estructura de las imágenes de entrenamiento. El modelo de apariencia que define la estructura de la imagen alrededor de cada marca se obtiene a partir de un perfil de píxeles muestreados en las direcciones perpendiculares al contorno. La dirección perpendicular a la marca  $(x_n, y_n)$  se obtiene rotando  $90^\circ$  el vector formado por las marcas  $(x_{n-1}, y_{n-1})$  y  $(x_{n+1}, y_{n+1})$ .

Una vez se han calculado las direcciones perpendiculares, se muestrean  $k$  píxeles a cada lado del contorno en cada una de las marcas para generar perfiles de longitud  $2*k+1$ . El modelo de apariencia se construye al obtener la primera derivada de cada uno de los perfiles, calculada como la diferencia entre los puntos  $j-1$  y  $j+1$ . Posteriormente, se normalizan los elementos del perfil de tal forma que la suma de sus valores absolutos sea 1.

A partir de  $N$  imágenes de entrenamiento se obtiene una serie de perfiles normalizados  $g_1, \dots, g_s$  para cada una de las marcas. Así, si se calculan el perfil medio  $g_m$  y la matriz de covarianzas  $S_g$ , es posible calcular la distancia de Mahalanobis entre un nuevo perfil y el modelo como:

$$f(g_i) = (g_i - g_m)^T * S_g^{-1} * (g_i - g_m) \quad (5)$$

## iii. Algoritmo de Búsqueda

El contorno de la región facial se ajusta mediante un proceso iterativo que comienza en el contorno medio. En cada iteración, cada marca se mueve  $n_s$  posiciones en la dirección perpendicular al contorno y se coloca en la posición que ofrece la menor distancia de Mahalanobis. Tras la actualización de las marcas, el modelo se desplaza a las nuevas posiciones.

A continuación, en la Figura 4, se muestra un ejemplo de la búsqueda de una cara utilizando el modelo explicado (Modelo de Forma Activa).



Figura 4. Búsqueda de una cara utilizando el Modelo de Forma Activa.

El modelo ASM es un buen método estadístico basado en la forma del objeto. Para ajustar un objeto determinado en una nueva imagen, se realizan deformaciones de forma iterativa. Las formas están restringidas por un modelo estadístico de forma, para variar sólo en las formas precisadas por un modelo de entrenamiento. Los puntos de referencia están localizados, generalmente, en los contornos. El gran problema de este modelo, es que son muy sensibles al proceso de inicialización, esto es, cuando la inicialización está alejada del objeto destino, ellos pueden converger hacia mínimos locales.

## 3. Bases de datos

### 3.1 Introducción

Durante la exposición de los experimentos desarrollados a lo largo del PFC, se han utilizado las siguientes bases de datos (Tabla 1):

Bases de datos	Marcado manual	Tipo de escenario	Número de sujetos
BioID	20 puntos	Controlado	23
Forensic ATVS	21 puntos	Controlado	50
MORPH	0 puntos	Semicontrolado	200 (seleccionados)
SCFace	21 puntos	No controlado	130

Tabla 1. Descripción bases de datos.

Las bases de datos MORPH y SCFace, empleadas en los experimentos 2 y 3, han sido marcadas, además, de forma automática.

### 3.2 BioID

La base de datos BioID, se puede obtener de forma gratuita en internet, ha sido creada y publicada para dar la posibilidad a todos los investigadores que trabajan en el área del reconocimiento facial, de comparar la calidad de sus algoritmos de detección de cara y puntos característicos faciales, con otros.

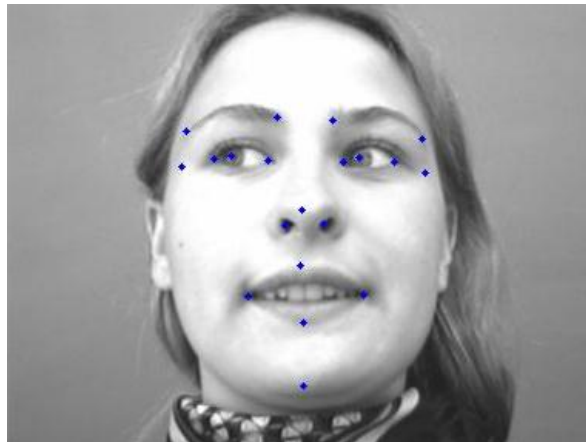
Incluye 1521 imágenes en escala de grises, todas ellas de un tamaño de 384x286 píxeles. Esta base de datos contiene imágenes de caras de 23 sujetos distintos, con una gran variedad de condiciones de iluminación, tamaños de cara, expresiones y fondos complejos. La Figura 5 muestra algunos ejemplos de dicha base de datos.



Figura 5. Cuatro usuarios, a modo de ejemplo, de la base de datos BioID

La base de datos BioID, ha sido marcada por David Cristinacce y Kola Babalola, estudiantes de Doctorado en el departamento de *Imaging Science and Biomedical Engineering (ISBE)*, en la Universidad de Manchester, de forma manual.

Se seleccionaron estos puntos de características faciales (Figura 6), debido a que son útiles para el análisis facial y reconocimiento de gestos.



- 0 = right eye pupil
- 1 = left eye pupil
- 2 = right mouth corner
- 3 = left mouth corner
- 4 = outer end of right eye brow
- 5 = inner end of right eye brow
- 6 = inner end of left eye brow
- 7 = outer end of left eye brow
- 8 = right temple
- 9 = outer corner of right eye
- 10 = inner corner of right eye
- 11 = inner corner of left eye
- 12 = outer corner of left eye
- 13 = left temple
- 14 = tip of nose
- 15 = right nostril
- 16 = left nostril
- 17 = centre point on outer edge of upper lip
- 18 = centre point on outer edge of lower lip
- 19 = tip of chin

Figura 6. Marcado de puntos BioID

### 3.3 Forensic ATVS

Otra base de datos utilizada en el proyecto, que ha sido proporcionada por el laboratorio ATVS, es la llamada *Forensic ATVS*, compuesta por imágenes adquiridas mediante una cámara fotográfica digital de gran calidad.

Las imágenes han sido tomadas en un periodo de tiempo de 1 año (2013), con 2 sesiones por cada usuario (entre las cuales pasaron, de media, 3 meses entre la primera sesión y la segunda) y 4 secuencias por sesión. Además, los usuarios fueron colocados dentro de una habitación con una iluminación estándar y un flash adecuado (condiciones estándar de las imágenes de pasaportes u otros documentos de identificación personal (por ejemplo, DNI)), en tres puntos distintos, seleccionados con anterioridad, a distintas distancias (1 metro, 2 metros y 3 metros). La figura siguiente, Figura 7, muestra la colocación de los usuarios:

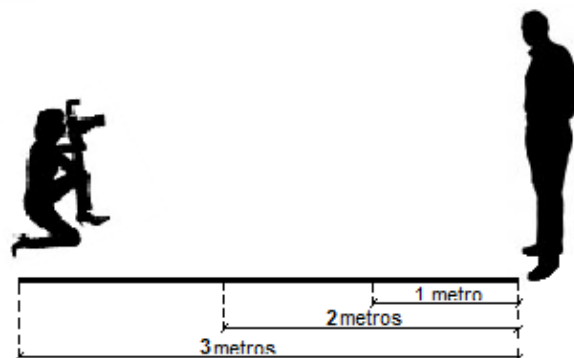


Figura 7. Distancias de captura en la base de datos ATVS.

La finalidad de sacar fotografías a tres distancias distintas, es la de ver el potencial de los algoritmos de reconocimiento facial para distintas calidades de imágenes. El cambio de calidad de las imágenes se puede observar en la Figura 8.

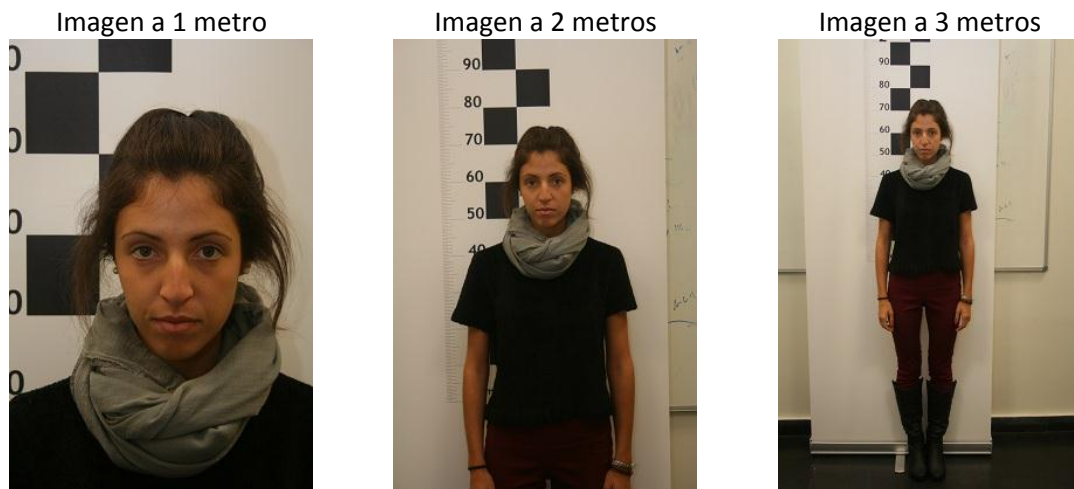
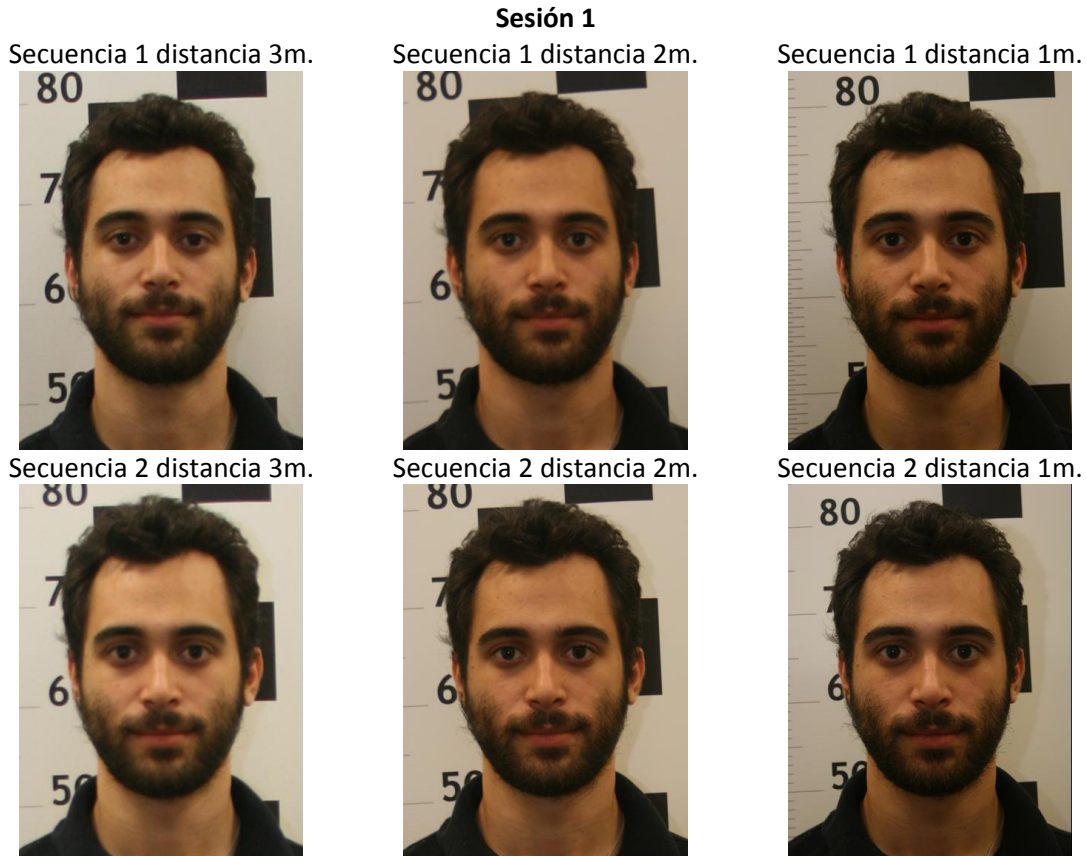


Figura 8. Imágenes ejemplo de la base de datos ATVS a diferentes distancias.

Los participantes de la base de datos, fueron estudiantes, profesores o empleados de la Universidad Autónoma de Madrid. Del total de los 50 voluntarios, 32 eran hombres y 18 mujeres. Todos los participantes fueron caucásicos, con un rango de edad de entre 21 y 65 años.

A continuación (Figura 9), se muestra el conjunto de imágenes, para un único usuario, disponibles en esta base de datos.





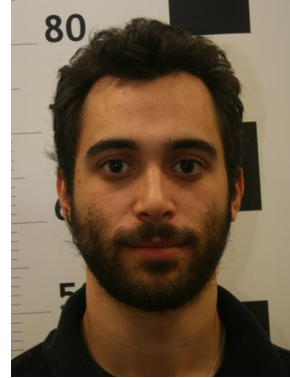
Secuencia 3 distancia 3m.



Secuencia 3 distancia 2m.



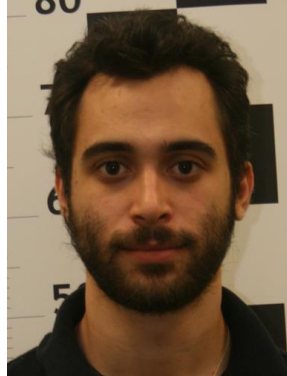
Secuencia 3 distancia 1m.



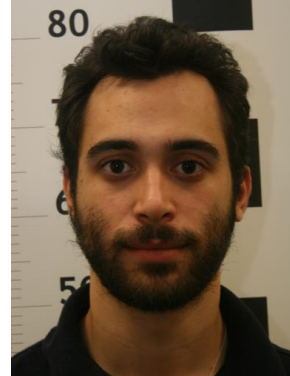
Secuencia 4 distancia 3m.



Secuencia 4 distancia 2m.



Secuencia 4 distancia 1m.

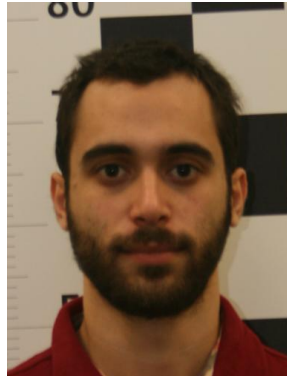


**Sesión 2**

Secuencia 1 distancia 3m.



Secuencia 1 distancia 2m.



Secuencia 1 distancia 1m.



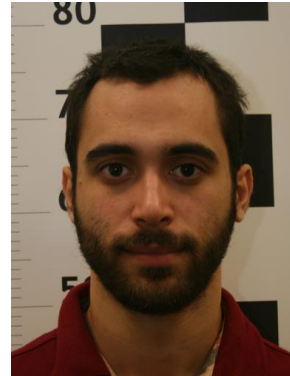
Secuencia 2 distancia 3m.



Secuencia 2 distancia 2m.



Secuencia 2 distancia 1m.





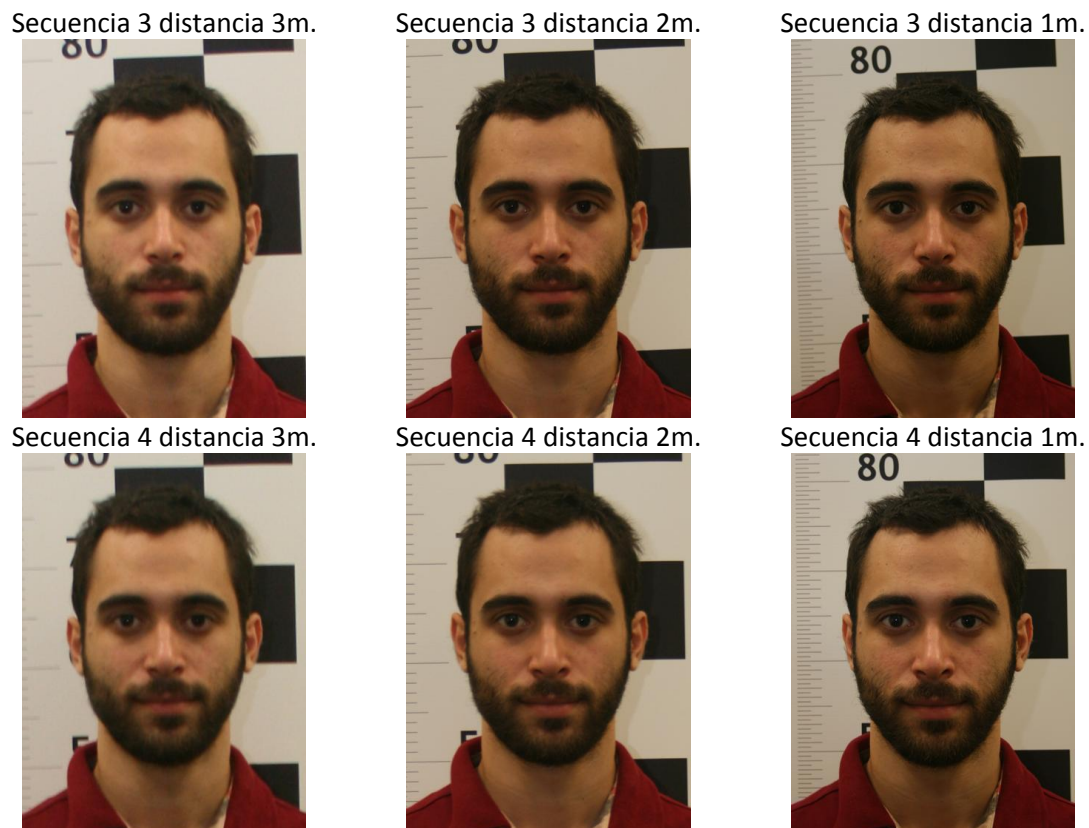


Figura 9. Conjunto de imágenes para un único usuario.

En esta base de datos, se proporcionan 23 puntos faciales (Figura 10), marcados de forma manual por una persona no experta.



Figura 10. Puntos faciales Forensic ATVS.

### 3.4 MORPH

La base de datos MORPH, versión no comercial, se recogió en un periodo de tiempo de 5 años (desde el año 2003 al 2007), con numerosas imágenes (concretamente, 55134 imágenes). Está

formada por más de 13000 individuos, con un rango de edad que se encuentra entre los 16 y los 77 años, con una edad media de 33 años.

El número de imágenes por individuo es variable, siendo el número medio de imágenes de 4. Esta base de datos, se trata de una colección adquirida en condiciones no controladas, es decir, fue recogida en condiciones del mundo real (entorno semi-controlado). Algunas de las imágenes de dicha base de datos se muestran en la Figura 11.



Figura 11. Cuatro usuarios, a modo de ejemplo, de la base de datos MORPH

La Tabla 2 muestra la distribución de imágenes por género y ascendencia.

	African	European	Asian	Hispanic	“Other”	Total
Male	36832	7961	141	1667	44	46645
Female	5757	2598	13	102	19	8489
Total	42589	10559	154	1769	63	55134

Tabla 2. Número de Imágenes por género y ascendencia (n=55134).

Cabe mencionar, que para el desarrollo del proyecto, solo se han utilizado las imágenes pertenecientes a la raza de Europeos.

La Tabla 3 muestra el número de imágenes adicionales que existen por usuario.

	1	2	3	4	5+	Total
Male	373	2350	3606	1975	3155	11459
Female	85	478	712	352	532	2159
Total	458	2828	4318	1308	3687	13618

Tabla 3. Número de imágenes por usuario.

La Tabla 4 muestra el número de imágenes faciales con respecto a la década de vida en que se hicieron las fotos.

	< 20	20 - 29	30 – 39	40 – 49	50+	Total
Male	6638	14016	12447	10062	3482	46645
Female	831	2309	2910	1988	451	8489
Total	7469	16325	15357	12050	3933	55134

Tabla 4. Número de Imágenes por género y edad (n=55134)

El conjunto de datos contiene metadatos en forma de edad, género, raza, altura, peso y los ojos (Tabla 5).

Subject Identifier (six digit ID with leading zeros)	Race ( B, W, A, H, O )
Picture Identifier ( beginning at 0 for each subject )	Gender ( M / F )
Date of Birth (mm/dd/yyyy)	Facial Hair Flag
Image Date (mm/dd/yyyy)	Glasses Flag
Age Difference ( number of days since last image )	
Image Filename ( ID_[picture_id][M/F][Age].JPG )	

Tabla 5. Detalles de los metadatos.

En las Figura 12, se presenta un conjunto de imágenes que demuestran la progresión de edad encontrada en este proyecto.



Figura 12. Imagen progresión edad hombre blanco (izq.) y mujer Afro-americana (dcha.)

En la base de datos MORPH, únicamente se proporcionan las coordenadas de los centro de los ojos.

### 3.5 SCFace

La base de datos SCface (surveillance cameras face database) fue diseñada principalmente como un medio para probar algoritmos de reconocimiento facial en condiciones del mundo real. En tal configuración, es fácil imaginar un escenario donde un individuo debería ser reconocido comparando una imagen frontal de la ficha policial con una obtenida de cámaras de vigilancia de calidad fija. Con el fin de lograr una configuración realista, se decidió utilizar cámaras de vigilancia comercialmente disponibles de calidad variable. La diferencia de esta base de datos con otras, es que las imágenes disponibles en las otras bases de datos, suelen ser obtenidas por la misma cámara y no se toman usando equipos de vigilancia comerciales.

SCface incluye imágenes que fueron tomadas en entornos interiores no controlados, usando cinco cámaras de vigilancia de calidad variable (Figura 13). Esta base de datos contiene 4160 imágenes (tanto en espectro visible, como de visión nocturna mediante infrarrojos (IR), ya que dos de las cámaras disponibles (cam1 y cam5) eran de visión nocturna) de 130 sujetos.

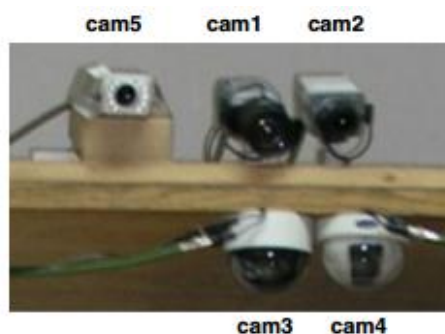


Figura 13. Posición cámaras comerciales de calidad variable.

La captura de las imágenes tuvo lugar en “*Video Communications Laboratory at the Faculty of Electrical Engineering and Computing*”, Universidad de Zagreb, Croatia [21]. El equipamiento utilizado fue: 6 cámaras de vigilancia de calidad variable (utilizadas para la adquisición de las imágenes no frontales), un ordenador (donde se almacenaron los videos e imágenes) y una cámara de fotos profesional de alta calidad (usada para la obtención de las imágenes frontales).

En la Figura 14, se observa el cambio de calidad tan ostentoso que hay entre las distintas cámaras.

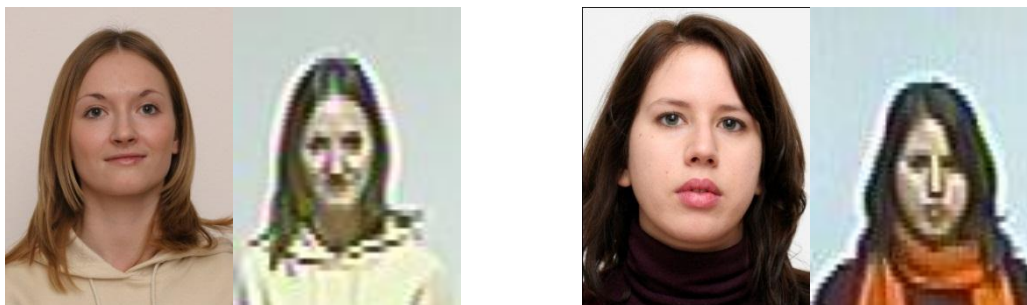


Figura 14. Imagen frontal (izquierda) e imagen camara1 (derecha) de dos usuarios de SCface.

Las 5 cámaras de vigilancia fueron instaladas en una habitación (cuya única fuente de luz era la luz exterior que entraba a través de la ventana situada en un lateral de la habitación, es decir, iluminación interior no controlada) de altura 2.25 metros y posicionadas como se muestra en la Figura 15.

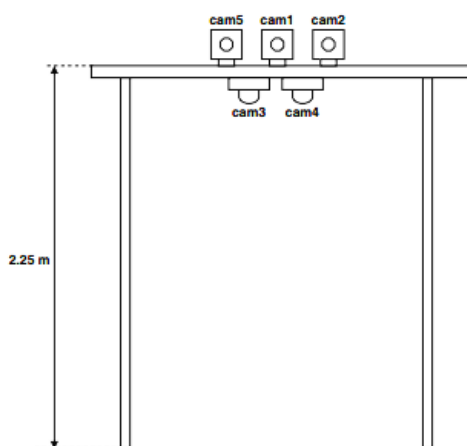


Figura 15. Cámaras de vigilancia.

Las características de las cámaras de vigilancia se recogen en la Tabla 6.

Type	cam1	cam2	cam3	cam4	cam5
	Bosch	Shany	J&S	Alarmcom	Shany
	LTC0495/51	WTC-8342	JCC-915D	VFD400-12B	MTC-L1438
CCD Type	1/3" IT	1/3" Sony Super HAD	1/3" Color	1/3" IT	1/3" Sony Super HAD
Active pixels	752×582	795×596	597×537	752×582	795×596
Resolution	540 TVL	480 TVL	350 TVL	460 TVL	480 TVL
Minimum illumination	0,24/0,038 lux (IR)	0,15 lux	0,3 lux	1,5 lux	0 lux (IR LED on at 4 lux)
SNR	> 50 dB	> 50 dB	> 48 dB	46 dB	> 50 dB
Video output	1 Vpp, 75 $\Omega$	1 Vpp, 75 $\Omega$	1 Vpp, 75 $\Omega$	1 Vpp, 75 $\Omega$	1 Vpp, 75 $\Omega$
Comment	IR night vision		dome camera	dome camera	IR night vision

Tabla 6. Especificaciones de las cámaras de vigilancia.

Para la adquisición de las imágenes de vigilancia, los usuarios tenían que pasear en frente de las cámaras e ir parándose en determinados puntos (Figura 16), previamente marcados, a distintas distancias (1 metro, 2.60 metros y 4.20 metros).

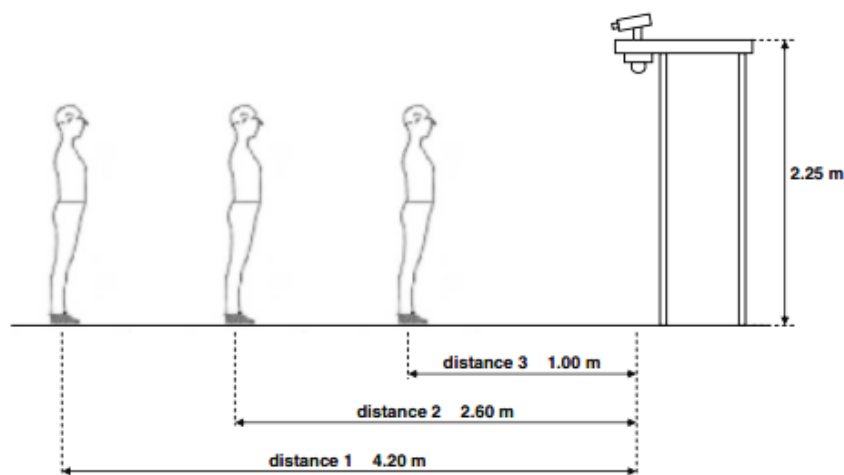


Figura 16. Distancias de captura de las imágenes de vigilancia.

En contraposición, las imágenes frontales, obtenidas mediante una cámara de gran calidad (especificaciones en Tabla 7), fueron hechas en una habitación separada, con una iluminación estándar y un flash adecuado (condiciones estándar de las imágenes de pasaportes u otros documentos de identificación personal (por ejemplo, DNI)).

Photographer's camera was Canon EOS 10D model with 22.7×15.1 mm CMOS sensor, with 6.3 mega pixels, equipped with Sigma 18–50 mm F3.5–5.6 DC lenses and Sigma EF 500 DG Super flash.

Tabla 7. Especificaciones cámara utilizada para imágenes frontales.



La captura de todas las imágenes, fue realizada en un periodo de tiempo de 5 días. A continuación (Figura 17), se muestra el conjunto de imágenes, para un único usuario, disponibles en esta base de datos.



Figura 17. Ejemplo del conjunto de imágenes por usuario.

Los participantes de la base de datos, fueron estudiantes, profesores o empleados de la Universidad, mencionada anteriormente. Del total de los 130 voluntarios, 114 eran hombres y 16 mujeres. Todos los participantes fueron caucásicos, con un rango de edad de entre 20 y 75 años. La distribución de edad se muestra en la Figura 18.

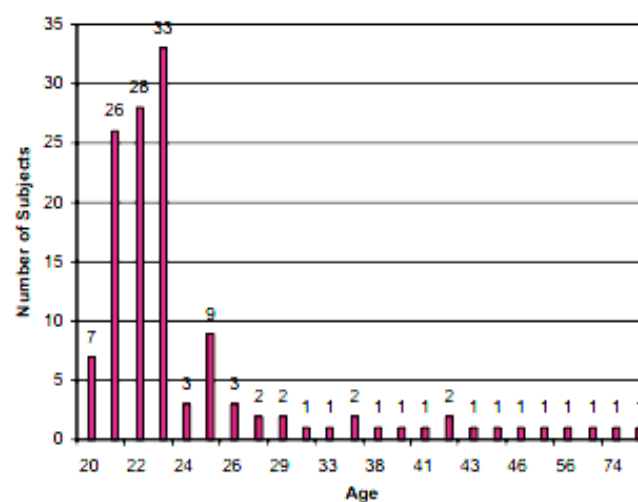


Figura 18. Distribución de la edad de los sujetos de la base de datos.

En la base de datos Scface, no se proporcionan ningún punto.

## 4. Sistema, diseño y desarrollo

### 4.1 Introducción a los sistemas biométricos

En la Figura 20 se muestra el diagrama de bloques general de un sistema de reconocimiento biométrico. En él pueden apreciarse tres módulos básicos: una *base de datos*, un *módulo de inscripción* y un *módulo de reconocimiento*. El conjunto de estos módulos realiza las funciones necesarias para reconocer a un individuo que accede al sistema.

- **Módulo de inscripción** (Figura 19). Este módulo está formado por un sistema de adquisición encargado de proporcionar la señal biométrica que caracteriza al individuo. Tras la adquisición de la señal biométrica se procede a la extracción de las características del rasgo biométrico del individuo. Dichas características expresan de una forma unívoca y compacta del individuo y constituyen su llamado *patrón biométrico*. El vector de características así formado, presenta menor dimensión que la señal previamente adquirida. De esta forma se realiza la codificación óptima de la señal en la que toda información irrelevante, que no contribuye al reconocimiento, es eliminada.

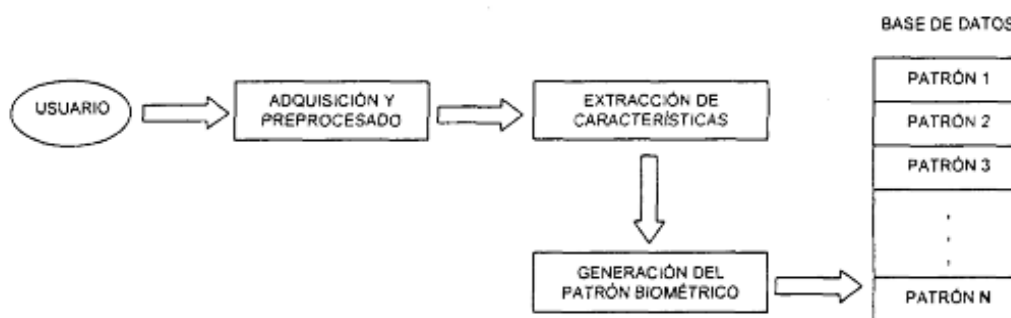


Figura 19. Módulo de inscripción.

- **Base de datos.** El patrón biométrico extraído por el módulo de inscripción es almacenado en la base de datos del sistema de reconocimiento. La base de datos contendrá, por tanto, todos los patrones biométricos de los individuos que sean usuarios legítimos del sistema.
- **Módulo de reconocimiento.** Este módulo se encarga de establecer la identidad del individuo que accede al sistema. Para ello, tras la adquisición del rasgo biométrico del individuo, se extraen las características y se obtiene el patrón biométrico, que, posteriormente, es comparado con los patrones almacenados en la base de datos. Los resultados de dichas comparaciones son cuantificados y valorados, permitiendo así la toma de decisiones respecto a la identidad del individuo en función de similitud obtenido.

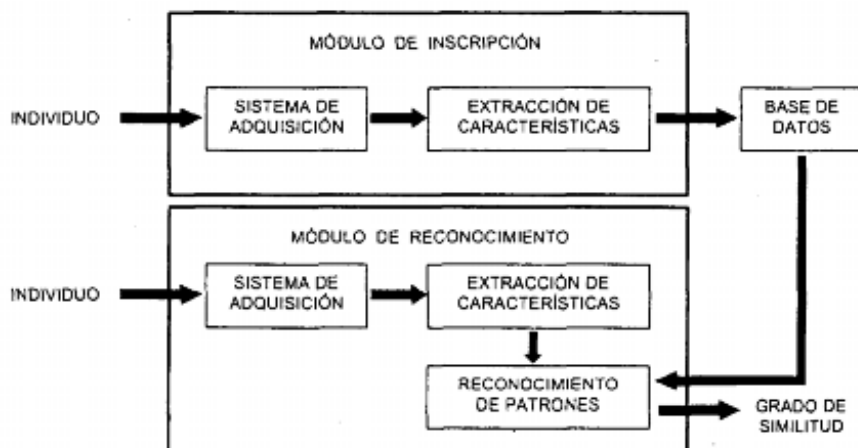


Figura 20. Diagrama de bloques de un sistema biométrico de reconocimiento.

Desde el punto de vista del funcionamiento de los sistemas automáticos de reconocimiento de personas mediante rasgos biométricos, se hace necesario clasificar las dos perspectivas fundamentales de trabajo de los mismos:

- Sistemas de reconocimiento en **modo identificación**.
- Sistemas de reconocimiento en **modo verificación**.

#### 4.1.1 Modo identificación

El objetivo es el de clasificar una realización determinada de un rasgo biométrico de identidad desconocida como perteneciente a uno de entre un conjunto de  $N$  posibles individuos. El esquema es el representado en la Figura 21.

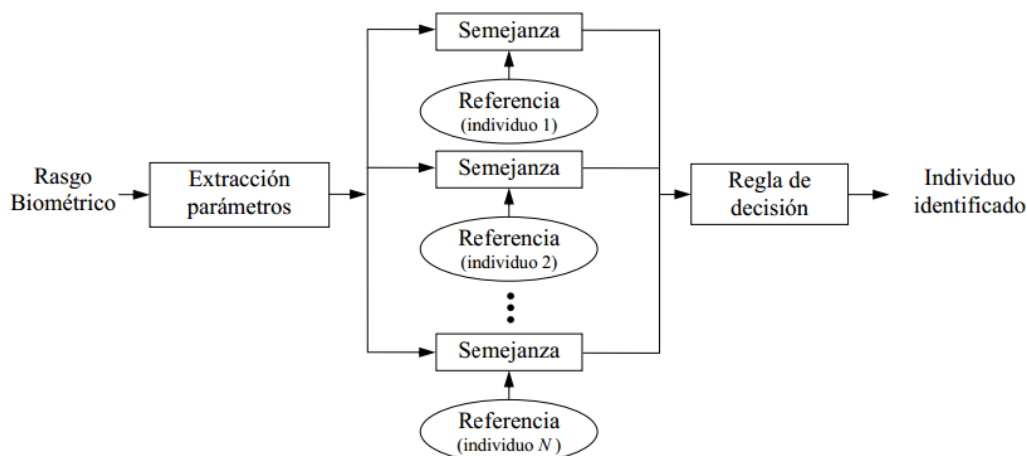


Figura 21. Esquema modo identificación.

Los errores que el sistema puede cometer durante el funcionamiento en modo de identificación son:

- El sistema se decide por el identificador de un patrón erróneo de la base de datos. Por tanto, el usuario no aparece como primer candidato de la lista.
- El usuario que demanda ser identificado, a pesar de aparecer como primer candidato en la lista, no obtiene el grado de semejanza suficiente como para ser reconocido como tal, y en consecuencia, el sistema considera que no pertenece a la base de datos.



Dentro de estos sistemas, debemos diferenciar dos posibles casos:

- **Identificación en conjunto cerrado:** en este caso, el resultado del proceso es una asignación de identidad a uno de los individuos modelados por el sistema, y conocidos como usuarios. Existen, por tanto,  $N$  probables decisiones de salida posibles.
- **Identificación en conjunto abierto:** aquí debemos considerar una posibilidad adicional a las  $N$  del caso anterior: que el individuo que pretende ser identificado no pertenezca al grupo de usuarios, con lo que el sistema de identificación debería contemplar la posibilidad de no clasificar la realización de entrada como perteneciente a las  $N$  posibilidades.

#### 4.1.2 Modo verificación

Los sistemas de verificación de individuos, por el contrario, toman dos entradas:

- una **realización del rasgo** biométrico a verificar,
- una **solicitud de identidad**, que puede ser realizada de diversas formas (lectura de tarjeta magnética individual, introducción mediante teclado o mediante voz de un código de locutor, etc.).

De este modo, las dos únicas salidas o decisiones del sistema son la **aceptación o rechazo** del individuo como aquél que pretende ser. De esta forma, el locutor solicitante será catalogado como **usuario auténtico** o bien como **impostor**, respectivamente.

La decisión de aceptar o rechazar la locución de entrada como correspondiente al locutor solicitado dependerá de si el valor de parecido o probabilidad obtenido supera o no un determinado **umbral** de decisión.

La figura siguiente (Figura 22) muestra de forma genérica la estructura típica de un sistema de verificación:

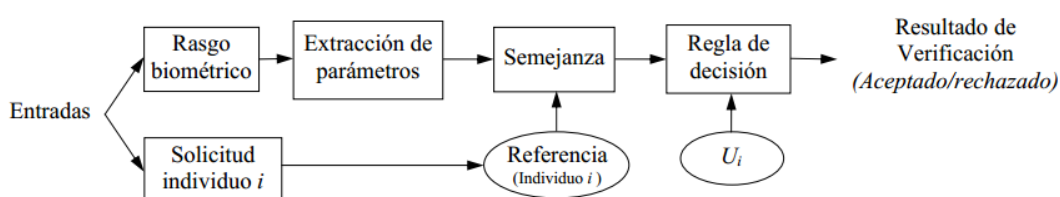


Figura 22. Esquema modo verificación.

En este caso, los errores que el sistema puede cometer durante el funcionamiento en modo verificación son:

- *Error de Falsa Aceptación (FA):* el sistema decide que el patrón almacenado corresponde a la identidad declarada por el patrón de entrada, cuando en realidad no lo es, aceptando a un individuo impostor.
- *Error de Falso Rechazo (FR):* en este caso, el sistema decide que el patrón almacenado no corresponde a la identidad declarada por el patrón de entrada, cuando en realidad sí lo es, rechazando al usuario legítimo.

### 4.1.3 Otras consideraciones

Los sistemas de verificación pueden ser vistos como un caso particular de identificación en conjunto abierto, en el que  $N=1$ .

Por otro lado, el proceso de identificación en conjunto abierto poder ser observado como un esquema en dos etapas:

- En la primera, se realizará un proceso de identificación en conjunto cerrado sobre los individuos conocidos, de forma que la realización incógnita sería asignada a uno de los usuarios del sistema con máxima probabilidad, con  $N$  decisiones de salida posibles.
- A continuación, se realizará un proceso de verificación respecto al individuo seleccionado en el trámite de identificación. De esta forma, la salida del sistema sería la aceptación del individuo que realiza el intento como uno de los  $N$  usuarios del sistema, o bien el rechazo al tomarlo como presunto impostor al mismo, lo que da lugar a un total de  $N+1$  decisiones posibles.

**El sistema de reconocimiento utilizado para los experimentos es en modo verificación, en conjunto cerrado.**

### 4.2 Sistema automático comercial

El sistema automático comercial utilizado para el marcado de puntos e identificación de usuarios de forma automática se denomina Luxand FaceSDK 4.0 [22].

Este sistema funciona para todas las cámaras existentes en el mercado: webcams, cámaras digitales... Es capaz de detectar caras con ángulos de giro de hasta  $\pm 30$  grados.

El SDK procesa imágenes, detectando la cara y mostrando como resultado 66 puntos faciales incluyendo ojos, contorno de los ojos, contorno de labios, cejas, contorno de la nariz y el contorno de la cara (se muestra en la Figura 23).



Figura 23. Puntos faciales obtenidos a través del FaceSDK.

Dicho sistema SDK es tratado como una caja negra con dos parámetros de configuración: i) indica el rango en el que se encuentra la imagen (dimensionalmente hablando) ii) indica el umbral de precisión (desde 1 hasta 5) a la hora de localizar caras (con umbral=1, el sistema te marca como cara cualquier región que el crea que es una cara, sin embargo, con umbral=5, solo marca regiones en las que tiene mucha seguridad).

**Los valores utilizados son 256 (hasta 500x500 píxeles de tamaño de cara) y 3.**

### 4.3 Descripción del sistema desarrollado

Hoy en día, el estado del arte del marcado de puntos faciales (más concretamente, el marcado de los ojos), está muy desarrollado y completo; sin embargo, el de detección de puntos característicos de la cara mal marcados y su posterior corrección está aún incipiente, y, por tanto, ha habido pocos avances en este tema.

Por ello, el proyecto no se centra tanto en el marcado de los puntos faciales (cabe destacar, que el sistema desarrollado utiliza los puntos característicos de un sistema comercial como punto de partida), sino, más bien, en detectar que rasgos de la cara está mal marcado, para poder corregirlo.

Para la cumplimentación de los objetivos del presente proyecto, se ha seguido el siguiente esquema (Figura 24):

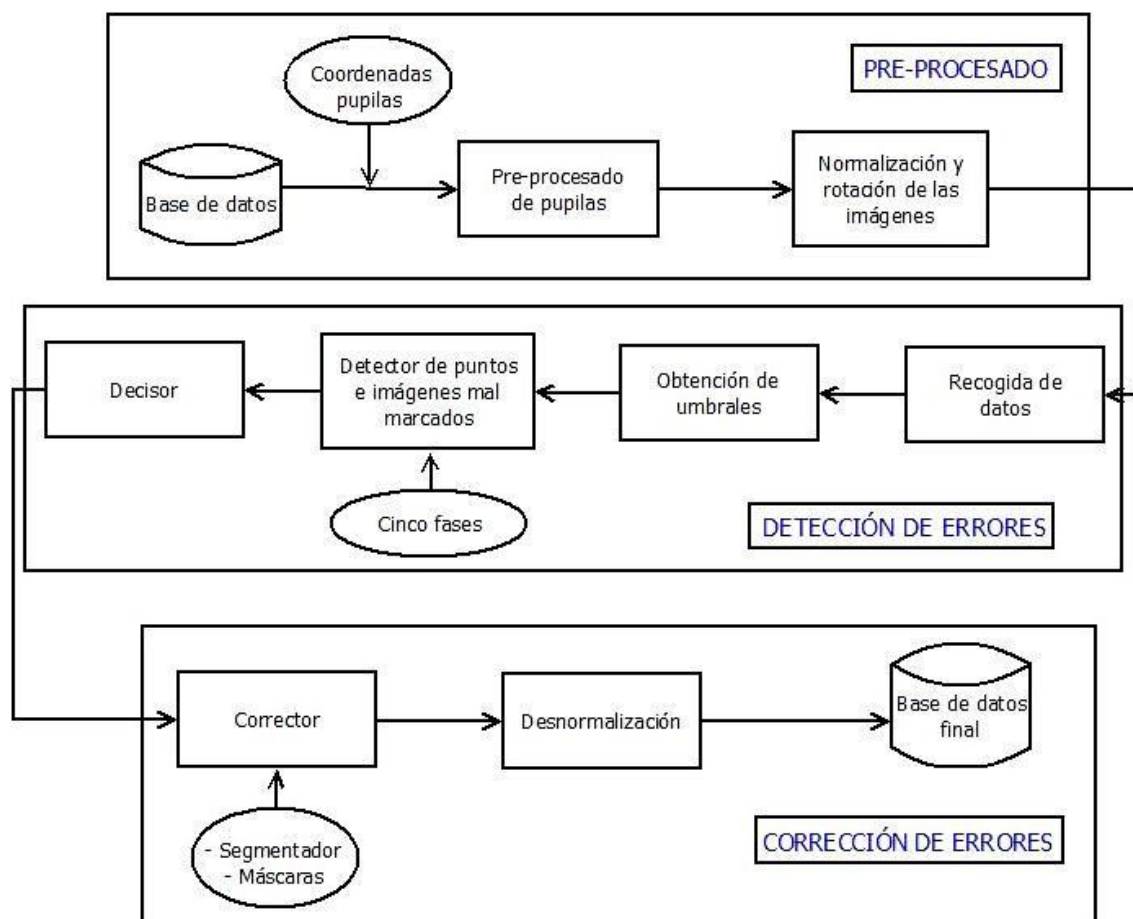


Figura 24. Esquema general para la detección y corrección de puntos característicos de la cara.

#### 4.4 Pre-procesado y normalización

En la actualidad, el marcado de los puntos característicos puede ser de manera manual, una persona realiza dicho marcado, o de manera automática, es decir, es un programa quien marca dichos puntos. En el mundo forense, el marcado utilizado es el manual, ya que tiene mayor precisión en estos momentos, pero en la realidad se tiende a automatizar el proceso, debido a que supone un ahorro de tiempo notable, de ahí el interés de este proyecto.

La adquisición de este sistema se ha realizado mediante cámaras fotográficas de calidad fija, en el caso de las bases de datos de BioID (entorno controlado) y MORPH (entorno semi-controlado). Sin embargo, en la base de datos de SCface, las imágenes han sido obtenidas por medio de cámaras de seguridad comerciales, con calidad variables (entorno no controlado).

En esta primera etapa, la calidad de la imagen capturada es una cuestión fundamental. El ruido, una mala regulación del contraste o iluminación de la cámara o una incorrecta posición del usuario (caras no frontales, objetos que ocluyan algún rasgo facial como las gafas de sol, bello facial, pelo, unas orejeras...) son algunos ejemplos que pueden ocasionar problemas en la tarea de reconocimiento.

Las imágenes de las caras utilizadas durante el diseño, desarrollo y experimentación del sistema implementado en este proyecto, se han obtenido de las bases de datos detalladas en capítulo 3.

Una vez adquiridas las imágenes, el primer paso, es el procesado de las imágenes (Figura 25)

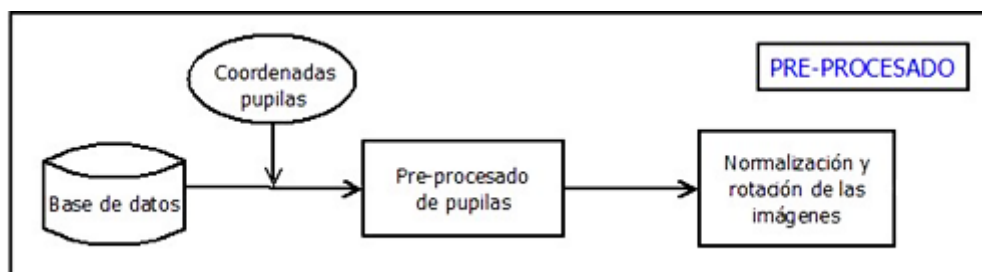


Figura 25. Etapa de pre-procesado.

Para la realización del sistema propuesto, además de contar con las imágenes, también se dispone de las coordenadas de las pupilas de los usuarios, ya que los sistemas automáticos generalmente se basan en el marcado de las pupilas. Con la finalidad de mejorar dichas coordenadas, se realiza un pre-procesado de las pupilas.

El objetivo del **pre-procesado de las pupilas** consiste en la obtención del punto medio del ojo, es decir, como las caras son frontales, se considera que todos los usuarios van a estar mirando hacia el frente, mejorando, además, la colocación de las pupilas de los usuarios (en la Figura 26 se muestra un ejemplo). Este método se efectúa para poder realizar de forma efectiva la normalización, alineación y rotación de las imágenes.

El nuevo valor de las pupilas se obtiene determinando el punto medio de los extremos de los ojos, consiguiéndose esto de la siguiente forma:

$$Pupila_{dcha_x} = \frac{(Punto_{exterior_ojo_{dchx}} + Punto_{interior_ojo_{dchx}})}{2}$$

$$Pupila_{dcha_y} = \frac{(Punto_{exterior_ojo_{dchy}} + Punto_{interior_ojo_{dchy}})}{2}$$

$$Pupila_{izqx} = \frac{(Punto_{exterior_ojo_{izqx}} + Punto_{interior_ojo_{izqx}})}{2}$$

$$Pupila_{izqy} = \frac{(Punto_{exterior_ojo_{izqy}} + Punto_{interior_ojo_{izqy}})}{2}$$

Este pre-procesado se aplica automáticamente cuando se dispone de los extremos de los ojos. En caso de que esto no suceda, es decir, el programa comercial no ha sido capaz de localizarlos, el sistema desarrollado en el proyecto, localiza dichos puntos de forma automática, en los casos en los que sólo se disponga de los puntos característicos de las pupilas. En la Figura 26 se muestra un ejemplo del pre-procesado de pupilas.



Figura 26. Marcado de las pupilas original (izquierda) y después del pre-procesado (derecha).

En la Figura 26, se aprecian las dos razones por el que es tan importante el pre-procesado de las pupilas: i) mal marcado de pupilas y ii) pupilas no situadas en el centro de los ojos. En las imágenes pertenecientes a la base de datos MORPH, se observa como el marcado de las pupilas (cruz verde) proporcionado por el sistema comercial Luxand es erróneo (foto izquierda), a pesar de estar mirando al frente el usuario, y, después de aplicar el algoritmo mencionado anteriormente de pre-procesado, las pupilas (puntos rojos) pasan a estar colocadas en una posición más exacta (imagen derecha), permitiendo una normalización y una rotación posterior de la imagen, más idónea.

En las imágenes proporcionadas por la base de datos Forensic ATVS (Figura 26), se observa como el usuario está mirando hacia uno de los lados, situándose así, las pupilas en uno de los lados de los ojos. Al realizar el pre-procesado, se consigue obtener el centro de los ojos, permitiendo de esta manera, una segmentación de la cara correcta.

Una vez hecho el pre-procesado de las pupilas, se realiza la normalización de imágenes, etapa clave con el objetivo de poder universalizar los algoritmos desarrollados en este proyecto, explicados más adelante.

La **normalización** se basa en el estándar ISO de imágenes [23]. A continuación, en la Tabla 8, se muestran las características geométricas que deben tener el tipo de imágenes utilizadas, así como un ejemplo gráfico (Figura 27).

Características o parámetros	Valor
Ancho de la imagen	$W$
Alto de la imagen	$W/0.75$
Coordenada Y de los ojos	$0.6*W$
Coordenada X del primer ojo (derecho)	$(0.375*W)-1$
Coordenada X del segundo ojo (izquierdo)	$(0.625*W)-1$
Ancho desde un ojo a otro (incluidos)	$0.25*W$

Tabla 8. Características geométricas de las imágenes.

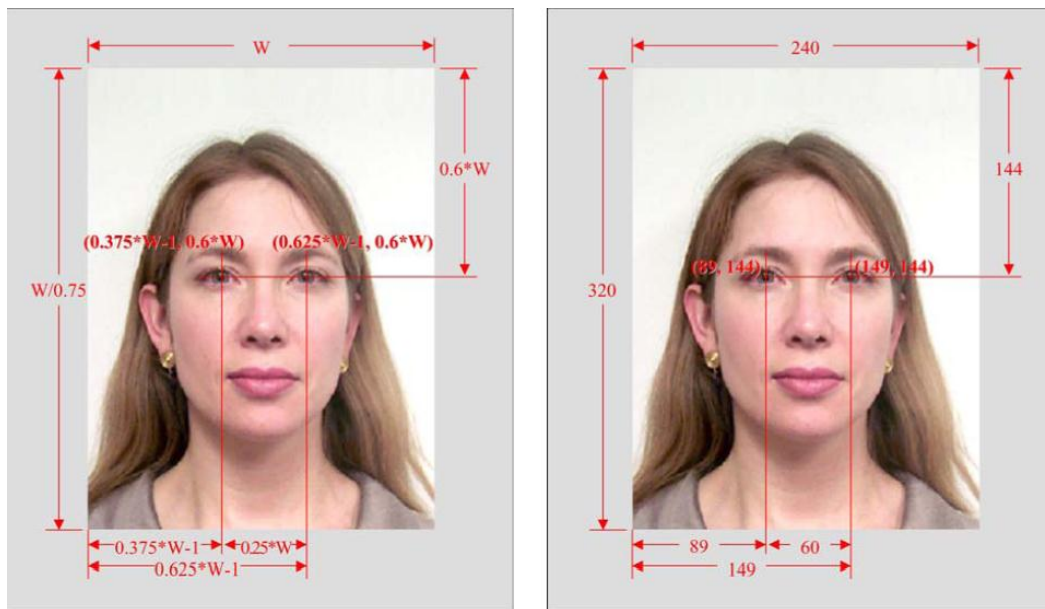


Figura 27. Formato geométrico (izquierda) y un ejemplo con  $W=240$  píxeles (derecha)

La normalización consta de tres partes: i) compensación de *rotación* con transformación afín, ii) compensación de *escalado*, es decir, todas las caras deben tener el mismo tamaño y la misma IPD, y, por último iii) *alineamiento*. Finalmente, se procede al recorte a partir de  $(x_0, y_0)$ , definido por el estándar ISO.

Por tanto, el objetivo de la normalización, no es solo para que todas las imágenes tengan las mismas dimensiones, sino para que las pupilas se encuentren situadas siempre en la misma posición, dentro de la imagen. Esto se consigue forzando que la distancia interpupilar (referenciado desde ahora por sus siglas en inglés IPD), es decir, la distancia entre los centros de los ojos, en nuestro caso, sea la misma para todos los usuarios de la base de datos.



Se han realizado pruebas con un ancho de imagen de  $W=168$  píxeles y  $W=300$  píxeles, eligiendo el valor de  $W=300$  píxeles para mantener una  $IPD=75$  píxeles, ya que es la más utilizada por la mayoría de sistemas automáticos de marcados de puntos característicos de la cara, asegurando siempre, una buena calidad.

El factor de dimensionado (ratio) utilizado para la obtención del nuevo tamaño de la imagen, se calcula mediante la siguiente ecuación:

$$ratio = \frac{IPD}{distancia\ real\ entre\ pupilas}$$

Como se observa en la Figura 28, la **rotación** se utiliza para enderezar una imagen en el plano frontal, tomando como eje de rotación la recta que pasa por los centros de los ojos de los usuarios. Esta transformación espacial, llamada transformación o aplicación afín, consiste en una transformación lineal seguida de una traslación o desplazamiento:

$$x \rightarrow Ax + b$$

Observando la fórmula anterior, se aprecia que esta transformación está compuesta por una matriz de rotación  $A$ , y un vector de traslación  $b$ , para ajustar las pupilas correspondientes a las coordenadas deseadas.

Para calcular la matriz de rotación se determina el ángulo  $\sigma$  entre la línea que une el centro de las pupilas de los ojos de las imágenes ( $d$ ) y la línea que une las coordenadas donde se desean que estén situados los centros de las pupilas ( $d_2$ ).

$$\alpha = \pm \arccos\left(\frac{d_2}{d}\right)$$

Con este ángulo (Figura 28), la transformada de rotación está definida como:

$$A = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) & 0 \\ -\sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Finalmente, el vector de traslación  $b$  permite superponer el punto que representa el centro de la pupila del ojo izquierdo, calculado mediante la rotación, y su equivalente en la imagen, y se determina hallando la distancia  $x$  e  $y$  entre ellos.



Figura 28. Imagen original (izquierda) e imagen normalizada y rotada (derecha).

Una vez obtenidas las nuevas imágenes, normalizadas y rotadas, se aplica dicha transformación matemática: rotación, alineación y escalado a todos los puntos característicos, de forma que queden normalizados sobre la imagen resultado.

En la Figura 29, se muestra un ejemplo con los pasos involucrados en una transformación de una imagen dada (explicado anteriormente), con el fin de la obtención de la cara. Para la creación de la imagen de ancho  $W=240$  píxeles, la imagen original (a) se gira para alinear horizontal los ojos (b). La imagen entonces, se escala uniformemente de modo que la distancia entre los ojos sea de 60 píxeles (c). Por último, la imagen se recorta (d) de manera que la coordenada del primer ojo (el derecho) sea de (89,144), es decir, 89 píxeles hacia la derecha y 144 píxeles hacia abajo desde la esquina superior izquierda de la imagen (0,0). Los píxeles negros, que son las fronteras del relleno, se pueden rellenar de cualquier otro color, siendo lo normal hacerlo del color que tengan los píxeles en la frontera de la imagen (e).



Figura 29. Transformación Afín y recorte.

#### 4.5 Detector de marcado incorrecto

Una vez superada la etapa de pre-procesado y realizada la normalización, tanto de los puntos característicos como de las imágenes de la base de datos, se comprueba la validez de los puntos característicos, marcado por un programa comercial automático (en nuestro caso es el llamado Luxand). Es decir, en este bloque (Figura 30) se comprueba la exactitud del marcado del resto de puntos característicos de la cara, rasgo a rasgo (pues se detecta el rasgo mal marcado), mediante el cálculo de distancias, ángulos y simetrías, pudiendo identificar aquellos rasgos faciales incorrectamente marcados.



Figura 30. Bloque detector



Para realizar dicha comprobación, el detector de puntos e imágenes mal marcadas (representado de manera esquemática en la Figura 31) consta de cinco fases: i) la primera fase, simplemente comprueba que la cara ha sido detectada y marcada, ii) esta segunda fase verifica si los centros de los ojos están posicionados en su lugar correcto (esta fase es muy importante y precisa de exactitud, ya que el resto del proyecto depende del correcto marcado de los centros de los ojos), iii) en esta fase se realiza un sistema de medición mediante distancias entre puntos, iv) la cuarta fase verifica si los ángulos que forman las cejas están dentro del rango definido, y v) esta quinta fase mide las simetrías básicas de la cara.

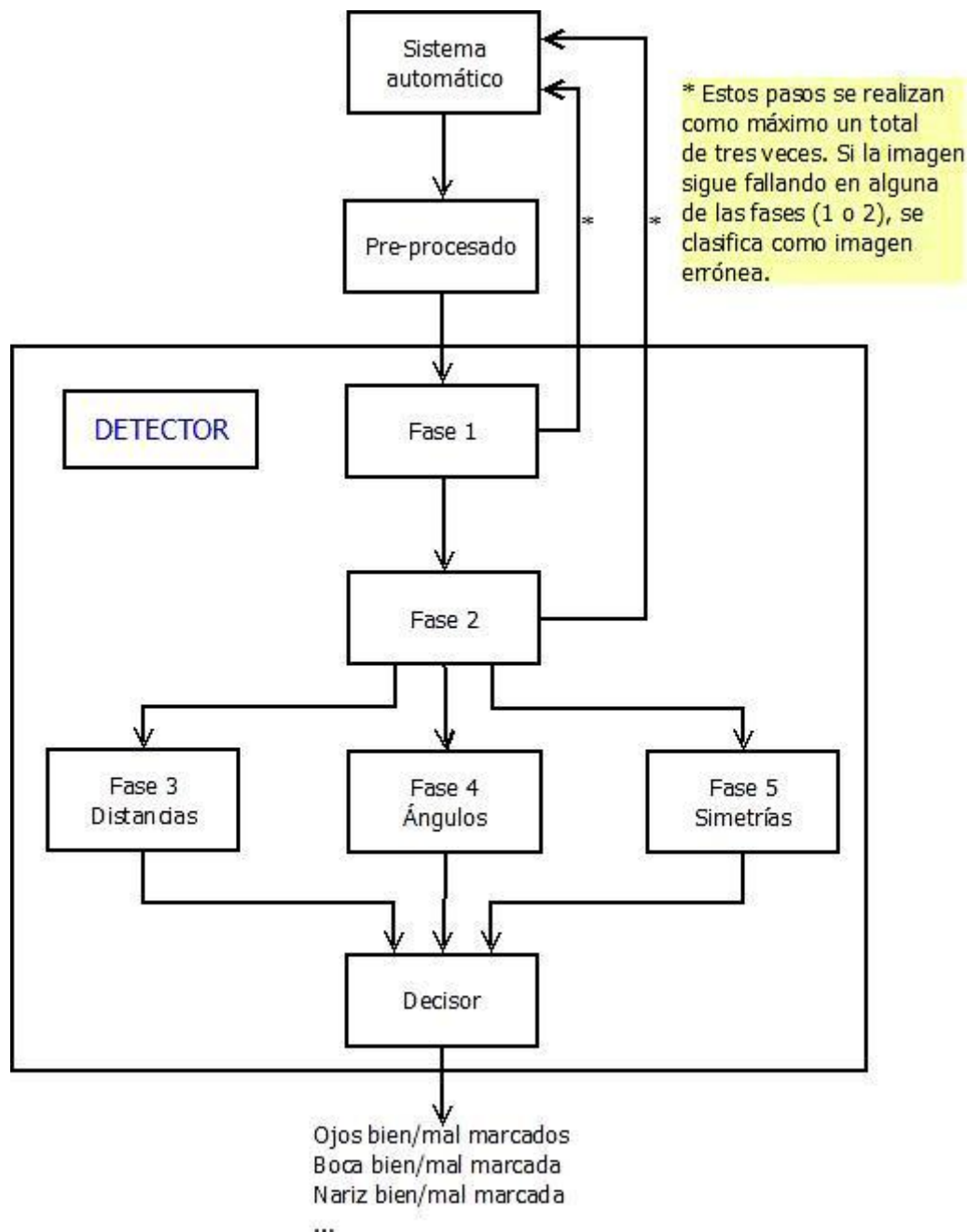


Figura 31. Detector de puntos característicos.

#### 4.5.1 Fase 1. No landmarks

En esta fase, simplemente se comprueba si el programa comercial FaceSDK de Luxand, ha sido capaz de encontrar los puntos característicos de la cara o no. En el caso de que no haya sido capaz, se pasa de nuevo al sistema para intentar, tuneando los parámetros de configuración, conseguir ajustar puntos. Si, repitiendo este proceso tres veces, no se consigue encontrar o ajustar bien los puntos, se deshecha la imagen de la base de datos, porque se considera que esa imagen es incorregible.

#### 4.5.2 Fase 2. Marcado incorrecto de ojos

Esta fase 2 es la más importante del detector, ya que todo el correcto funcionamiento del sistema desarrollado depende del correcto marcado de los centros de los ojos. Por ello, el objetivo de esta fase radica en detectar si el centro de los ojos ha sido correcta o incorrectamente marcado.

Para llegar a dicho objetivo, se detecta si los ojos (más concretamente, centro de los ojos) se encuentran, no solo dentro de la cara, si no en una zona no lisa de la cara, es decir, los mofletes, la frente... Esto se consigue mediante el procedimiento de dilatación. La motivación de utilizar la dilatación, es que tiene una baja complejidad computacional.

Los elementos estructurantes (2 dimensiones) utilizados en el proyecto se representan en la Figura 32:



Figura 32. Elementos estructurantes planos

Los valores oscuros representan ceros y los valores claros representan  $-\infty$ . El punto negro indica el origen de coordenadas.

El algoritmo utilizado en esta fase, consiste en la transformación de la imagen original, en color, en una imagen en escala de grises, para, posteriormente, calcular los bordes a través del método de Canny, binarizando, de esta forma, la imagen. Posteriormente, se realizan varias dilataciones, con los dos operadores morfológicos explicados anteriormente, consiguiendo unir los rasgos faciales entre sí. Finalmente, se realiza una comprobación, para determinar si las coordenadas de los centros de los ojos están situadas dentro o fuera de la cara, y si se encuentran en una zona de la cara sin arrugas, rasgos característicos cercanos... Esta comprobación consiste en mirar si la suma de la cantidad de píxeles en blanco, dentro de un vecindario, el cual consiste en una línea horizontal centrada en el centro de los ojos, de ancho/20 píxeles (para el caso de  $W=168$ , vecindario=9 píxeles), tiene como resultado, al menos, un treinta por ciento (30%) de píxeles en blanco.

A continuación, en la Figura 33, se representan imágenes (usuario de la base de datos MORPH) explicativas del procedimiento explicado con anterioridad, las cuales pasan la fase.

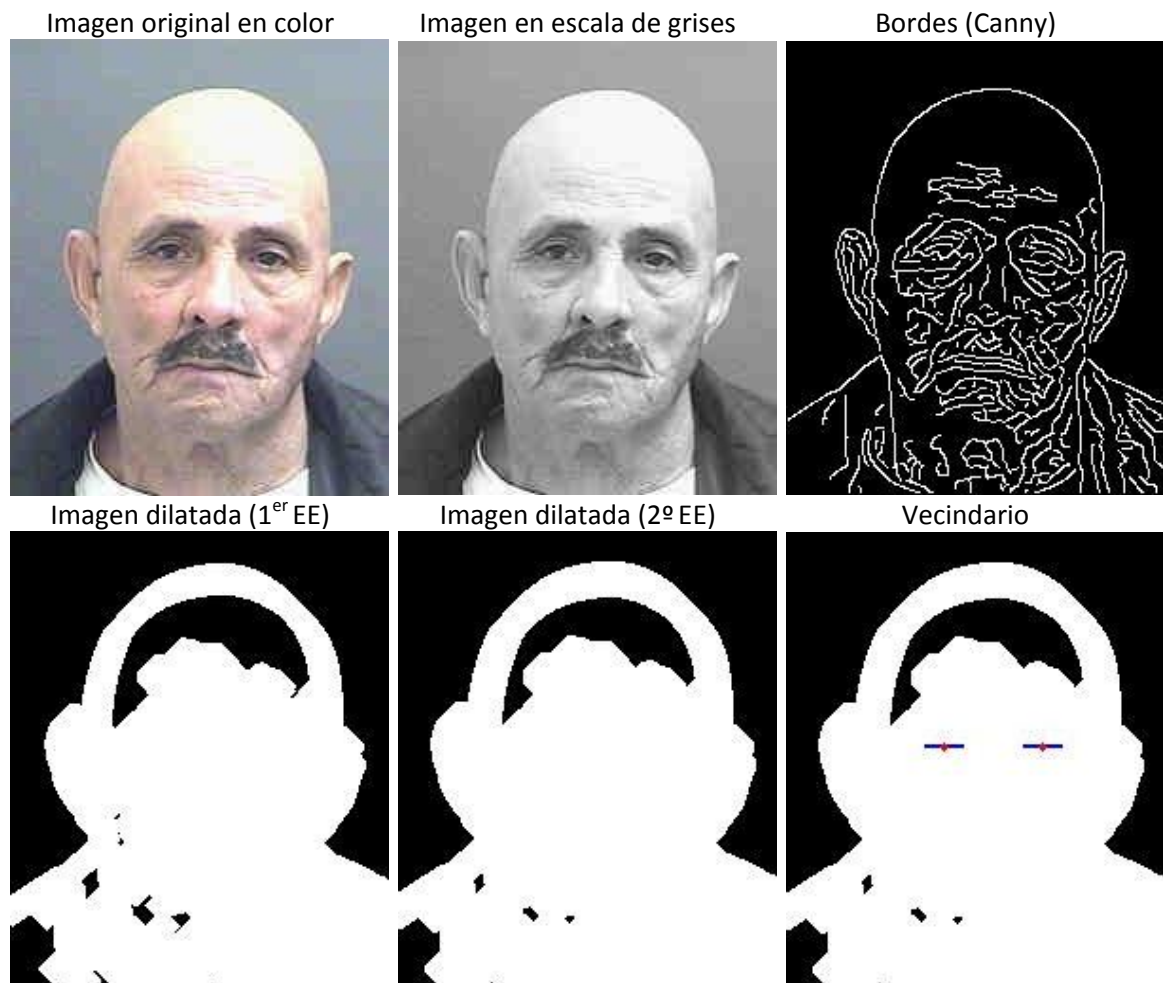


Figura 33. Representación fase 2

En la Figura 34, se muestra el caso particular en que la imagen no pasa esta fase, debido a que el centro de los ojos se encuentra en una parte de la cara lisa (en este caso, se encuentran sobre la frente).



Figura 34. Ejemplo imágenes con mal marcado de ojos (ojos en parte lisa de la cara).

Como se ha mencionado anteriormente, debido a la importancia que tiene el correcto marcado de los centros de los ojos, adicionalmente, se crea, dentro de esta fase 2, otra etapa (mucho más estricta) para ser capaces de detectar si las pupilas han sido bien o mal marcadas. Esta segunda etapa, solventa los errores de marcado no detectados por la etapa anterior (se grafica un ejemplo en la Figura 35).

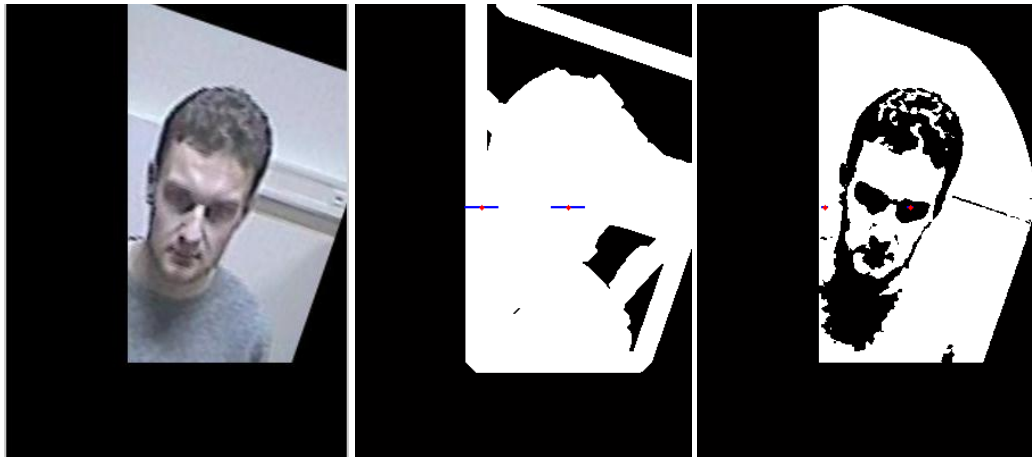


Figura 35. Restricciones de la etapa 2, fase 2.

La imagen central, en la Figura 35, ha pasado la primera etapa de esta segunda fase, debido a que la dilatación ha producido que se junten la cara del usuario con el fondo, pareciendo que los centros de los ojos han sido marcados de forma correcta, ya que, teóricamente, se encuentran dentro de la cara y sobre un rasgo facial. Sin embargo, como se observa en la imagen de la derecha, esto no es así, ya que el ojo derecho se encuentra fuera de la cara, lo cual es detectado por la segunda etapa.

Para la realización de esta segunda etapa, se aplica una máscara elíptica a la imagen, mostrando solo la región que nos interesa, en este caso, la cara. En caso de que la imagen sea a color, se transforma a escala de grises. Posteriormente, se realiza un realce de contraste, es decir, se aclara la imagen, convirtiendo los grises más oscuros, a otros más claros. Finalmente, se binariza la fotografía, tomando ésta, valores de 0 o 1 únicamente. A continuación, se suman la cantidad de píxeles en blanco, dentro de un vecindario de seis píxeles (zona de interés), alrededor de las coordenadas de las pupilas. Nuevamente, el vecindario es una línea horizontal, pero, como se puede observar en la Figura 36, muy restrictiva. Finalmente, se realiza una comprobación, para ver si las coordenadas de las pupilas están dentro o fuera de una zona de la cara sin arrugas, rasgos característicos cercanos, etc. Esta comprobación consiste en mirar si la suma realizada anteriormente, tiene como resultado, al menos, un píxel en blanco dentro del vecindario.

En la Figura 36, se representan imágenes explicativas del procedimiento explicado con anterioridad, que pasan la fase.

Imagen original en color



Imagen + máscara elíptica



Imagen en escala de grises





Figura 36. Ejemplo imagen marcada de forma correcta.

#### 4.5.3 Fase 3. Distancias

El objetivo de esta fase, es la medición de las distancias (distancia 1 y distancia 2) que hay entre el centro de la boca y cada una de las pupilas de los usuarios. Estos datos son tomados a partir de las imágenes normalizadas, y se almacenan en un vector, que se utilizará posteriormente para la obtención de los umbrales. En la Figura 37 se muestra un ejemplo de las distancias calculadas, que nos permiten identificar si el rasgo de la boca se encuentra incorrectamente marcado.



Figura 37. Distancias calculadas

#### 4.5.4 Fase 4. Ángulos

El objetivo de esta fase, es el cálculo de los ángulos que forman la recta que une el centro de ambas cejas (punto central de estas) y la recta que une cada uno de los centros de los ojos de los usuarios. Estos datos son tomados a partir de las imágenes normalizadas, y se almacenan en un vector, que se utilizará posteriormente para la obtención de los umbrales. Para un mayor entendimiento, en la Figura 38 se muestra un ejemplo:



Figura 38. Ángulos.

#### 4.5.5 Fase 5. Simetrías

El objetivo de esta fase, es la comprobación de la simetría entre los puntos característicos de la cara. Esto se comprueba mediante la medición de las distancias que hay entre la recta que une ambas pupilas de los usuarios, recta horizontal, y cada uno de los puntos característicos de la

cara (medidas verticales) y las distancias que hay entre la recta que pasa, de forma vertical, por el punto medio entre las pupilas y cada uno de los rasgos característicos de la cara (medidas horizontales). Estos datos son tomados a partir de las imágenes normalizadas, y se almacenan en un vector, que se utilizará posteriormente para la obtención de los umbrales.

- Distancias a la vertical:
  - Ojos:
    - Desde el lado exterior del ojo derecho (D)
    - Desde el lado interior del ojo derecho (E)
    - Desde el lado exterior del ojo izquierdo (D')
    - Desde el lado interior del ojo izquierdo (E')
  - Cejas:
    - Desde el lado exterior de la ceja derecha (A)
    - Desde el lado interior de la ceja derecha (C)
    - Desde el punto central de la ceja derecha (B)
    - Desde el lado exterior de la ceja izquierda (A')
    - Desde el lado exterior de la ceja izquierda (B')
    - Desde el punto central de la ceja izquierda (C')
  - Nariz:
    - Desde el lado izquierdo de la nariz (F')
    - Desde el lado derecho de la nariz (F)
  - Boca:
    - Desde el lado derecho de la boca (G)
    - Desde el lado izquierdo de la boca (G')
  - Barbilla:
    - Desde el punto central de la barbilla (H)
- Distancias a la horizontal:
  - Ojos:
    - No tiene sentido (imágenes ya normalizadas  $\rightarrow (D \approx D' \approx E \approx E' \approx 0)$ )
  - Cejas:
    - Desde el lado exterior de la ceja derecha (A)
    - Desde el lado interior de la ceja derecha (C)
    - Desde el punto central de la ceja derecha (B)
    - Desde el lado exterior de la ceja izquierda (A')
    - Desde el lado exterior de la ceja izquierda (B')
    - Desde el punto central de la ceja izquierda (C')
  - Nariz:
    - Desde el lado izquierdo de la nariz (F')
    - Desde el lado derecho de la nariz (F)
  - Boca:
    - Desde el lado derecho de la boca (G)
    - Desde el lado izquierdo de la boca (G')
  - Barbilla:
    - Desde el punto central de la barbilla (H)

Seguidamente, se muestra una imagen (Figura 39) con las distancias marcadas.

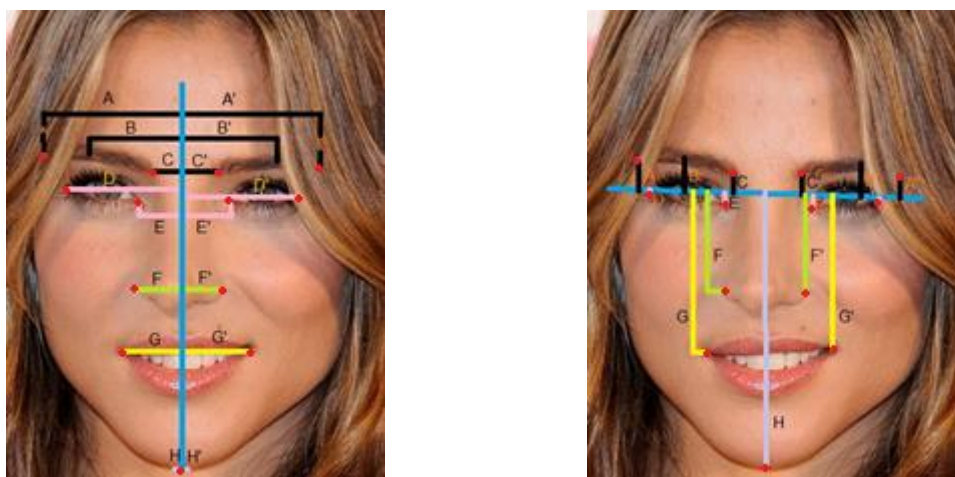


Figura 39. Distancias a la vertical (izquierda) y distancias a la horizontal (derecha).

Los umbrales necesarios para determinar si las medidas explicadas están dentro de un rango correcto (media de todas las medidas de un mismo tipo  $\pm$  umbral), se calculan, de forma automática, mediante una función, cuyos valores de entrada son el vector que contiene todas las medidas de un mismo tipo, por ejemplo, los ángulos  $\alpha$  de todos los usuarios registrados de la base de datos, y el porcentaje, es decir, el tanto por ciento que se quiere eliminar de todas las medidas, porque se estiman incorrectas. El conjunto de medidas tomadas siguen una distribución gaussiana, por tanto, el porcentaje se aplica a ambas colas de dicha distribución, eliminando los datos situados en las colas por debajo del valor elegido. Los valores de salida de esta función, son los umbrales, tanto el inferior como el superior.

El funcionamiento de dicha función, se basa en la obtención, y posterior normalización, del histograma de cada uno de los vectores de datos, para poder realizar una búsqueda de los valores del histograma normalizado que sean menores o iguales al porcentaje introducido; realizándose en ambas colas del histograma. En el caso de que el histograma solo tenga una cola (por ejemplo la derecha), al umbral perteneciente a la cola que no existe (la izquierda) se le asigna un valor muy alto para que no sea restrictivo.

En la Figura 40, se muestran los histogramas obtenidos para los ángulos  $\alpha$  de todos los usuarios de la base de datos de MORPH, únicamente raza europea, donde se puede ver que el rango en el que se encuentran la mayor parte de los ángulos  $\alpha$  de los usuarios registrados es  $\pm 5$  grados, aproximadamente. El signo que acompaña al ángulo  $\alpha$  (+,-), indica que ceja es la que está más arriba, siendo la ceja izquierda si  $\alpha$  es negativo, y la derecha en caso contrario ( $\alpha$  positivo).



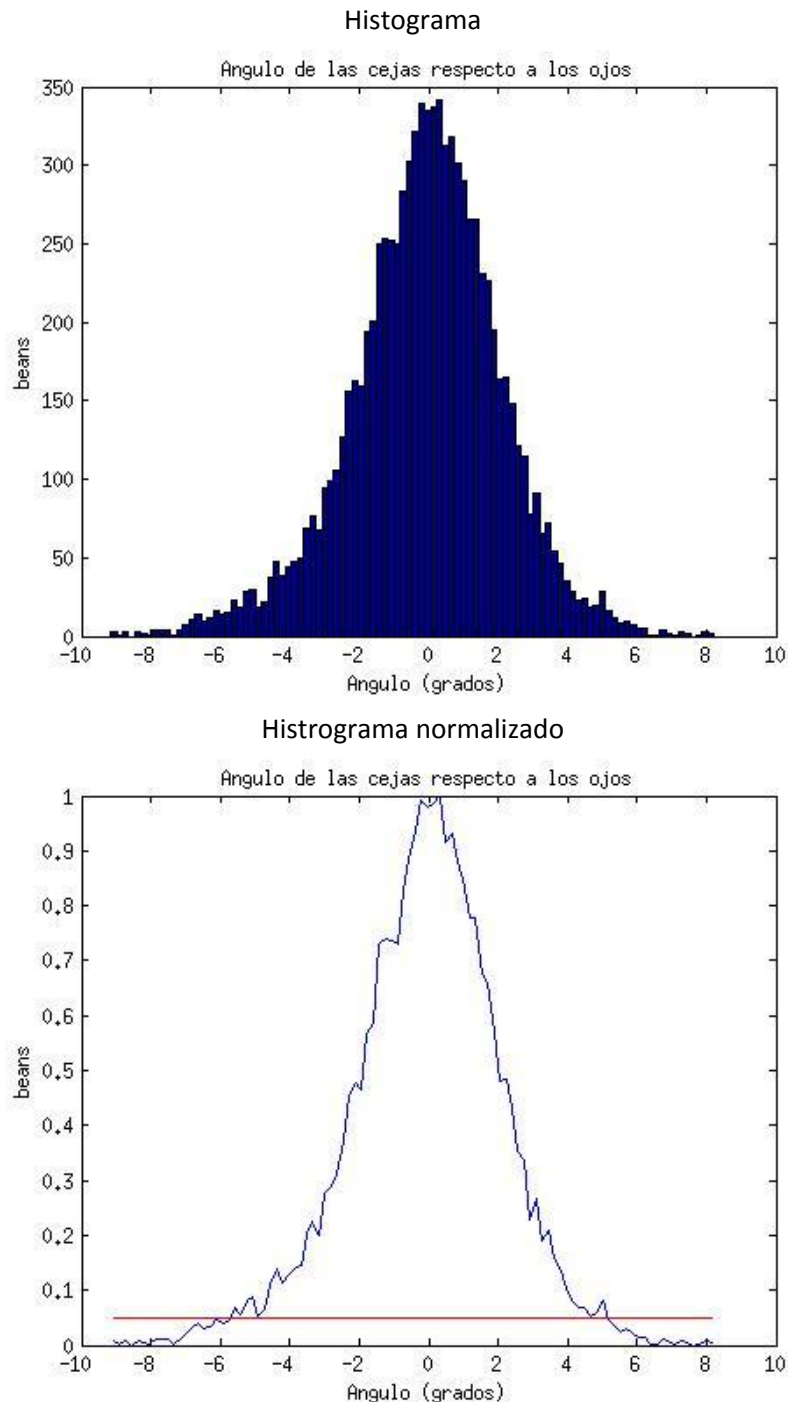


Figura 40. Histogramas resultado para ángulos  $\alpha$ .

La recta de color rojo representada en el histograma normalizado, indica la cota superior de los valores que son eliminados, es decir, los valores del histograma (azul) que se encuentran por debajo de la línea roja (porcentaje indicado como argumento de entrada, en este caso 5%), son eliminados porque se salen de los valores habituales.

A continuación, en la Figura 41, se muestran imágenes de la base de datos MORPH (europeos), que han fallado en alguna de las fases (en azul se muestran los puntos examinados, que dictan, mediante el umbral calculado anteriormente, si el punto característico está bien marcado).





Figura 41. Ejemplo de imágenes que fallan en cada una de las fases del detector, para la base de datos MORPH

#### 4.5.6 Decisor

Este bloque (Figura 42) elige, dentro de las imágenes que han sido clasificadas como erróneas, pero no obstante corregibles, es decir, las imágenes que no han pasado las fases tres, cuatro, y/o cinco, cual es el rasgo que se considera mal marcado, para poder centrarse, más concretamente, en la corrección de los puntos característico de ese único rasgo.

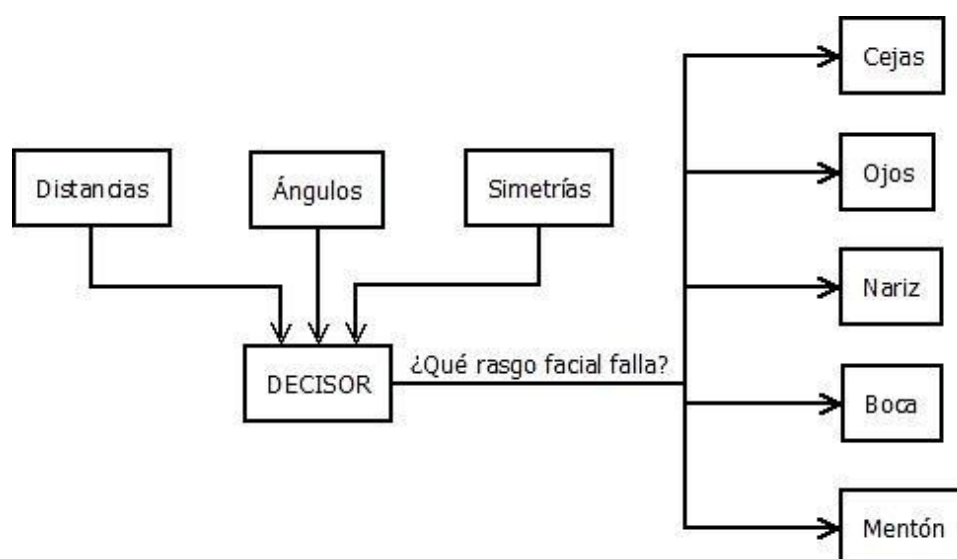


Figura 42. Esquema decisor.

Esta elección, se realiza de la siguiente forma:

1. Si la imagen errónea falla en la fase 3 (distancias), se considera que el rasgo mal marcado es la boca.
2. Si la imagen errónea falla en la fase 4 (ángulos), se considera que los rasgos mal marcados son las ambas cejas.
3. En caso de que la imagen errónea falle en la fase 5 (simetrías), hay que analizar la simetría del rasgo que ha fallado.
  - Se considera, en esta fase, que un rasgo ha fallado si ambas simetrías (tanto la horizontal como la vertical) de dicho rasgo están fuera del umbral elegido.

Como tanto la fase tres, la cuatro y la cinco no son excluyentes, una imagen puede fallar en varios rasgos, o un mismo rasgo puede fallar en varias fases (por ejemplo, la boca puede fallar en la fase tres y en la fase cinco).

A continuación (Figura 43) se muestran imágenes ejemplo (de la base de datos MORPH), las cuales han fallado en el marcado de algún rasgo (puntos en azul), y una imagen sin fallos, para que se pueda realizar una comparativa:

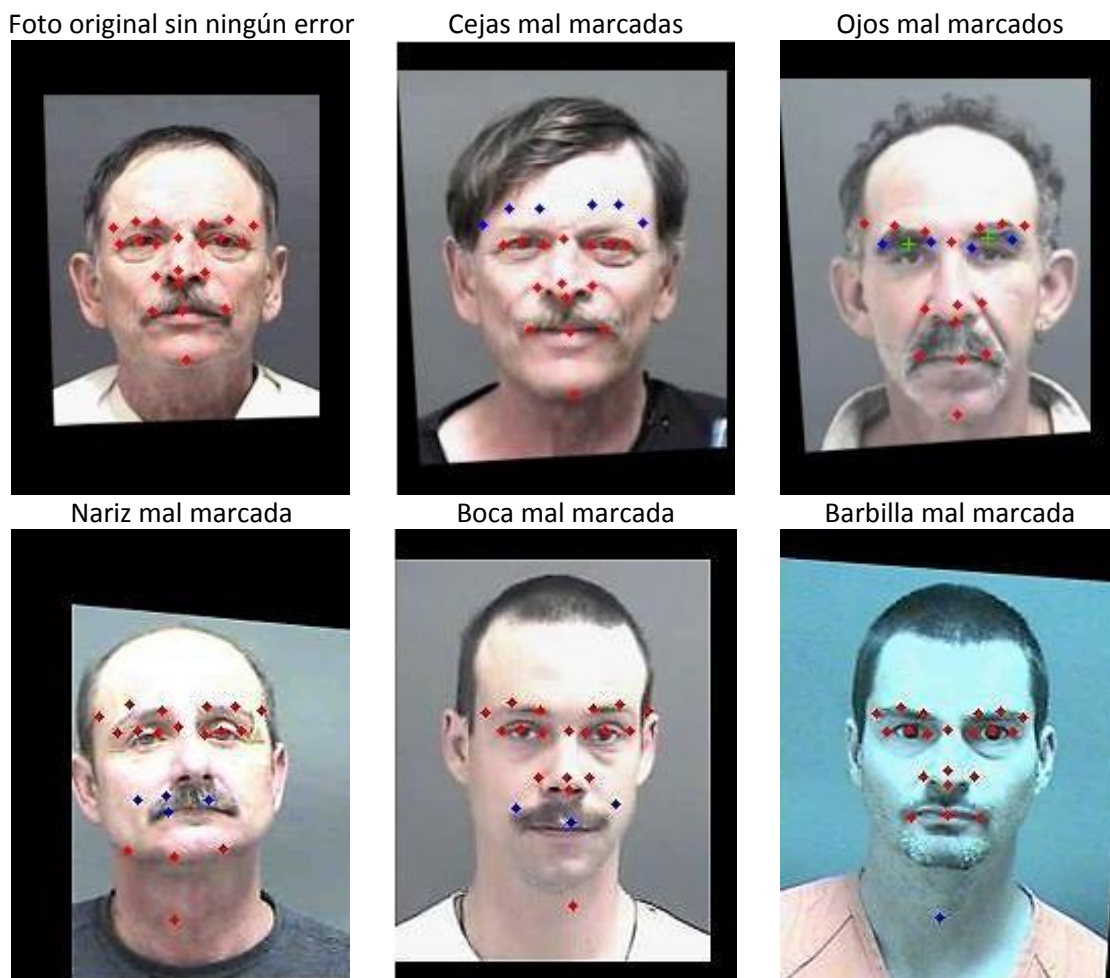


Figura 43. Imágenes mal marcadas de MORPH.

#### 4.6 Corrector de puntos característicos

Después de la fase de detección, las imágenes que se han clasificado como mal marcadas, pasan al corrector (Figura 44), mejorando de esta manera, el marcado automático de los puntos característicos de la cara.

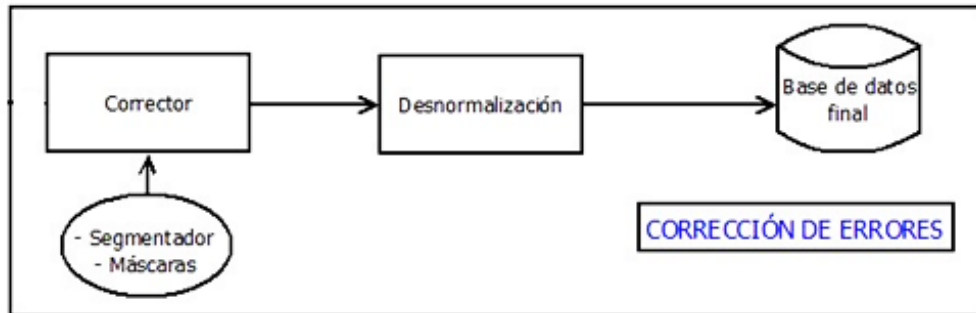


Figura 44. Bloque corrector de puntos faciales.

Dicha corrección (Figura 45), no se realiza en todos los puntos característicos de la cara (lo que sería una pérdida de recursos y de tiempo) sino, exclusivamente, en los puntos de los rasgos faciales que el decisor ha sacado como erróneos. Para realizar la corrección rasgo a rasgo, lo ideal es segmentar la imagen (extrayendo los rasgos faciales de la cara), para acotar la región de interés y no marcar los puntos de un rasgo en toda la imagen, simplificando la tarea de marcado/corrección automática.

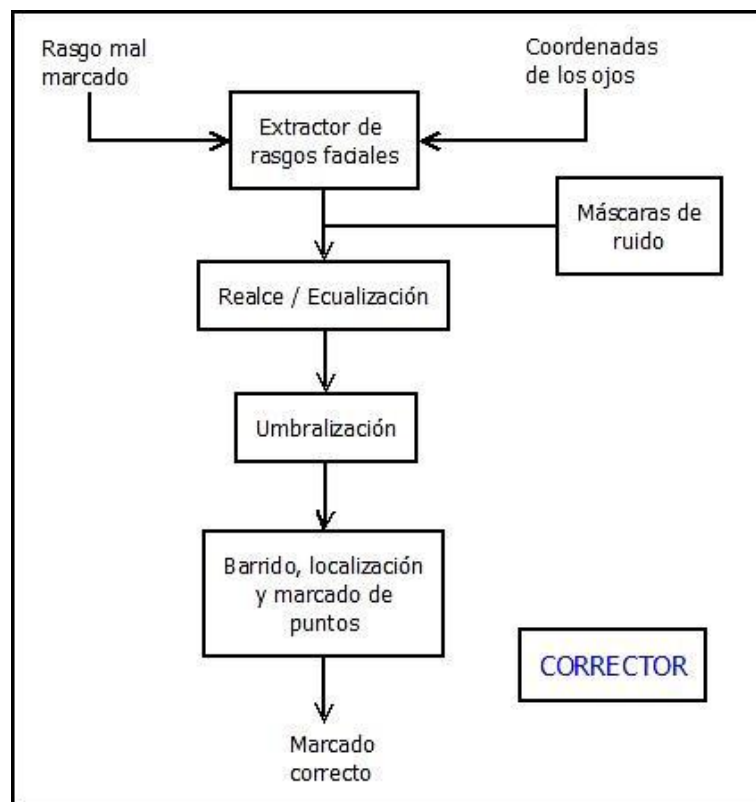


Figura 45. Funcionamiento corrector puntos característicos.

Para ello, se desarrollan dos *extractores de rasgos faciales*: i) uno que está basado en proporciones faciales, para entornos no controlados, y ii) otro que se basa en un punto medio para entornos controlados, o entornos no controlados, pero previo paso de los puntos marcados por un corrector automático.

#### 4.6.1 Extractor de rasgos faciales por proporciones

Como en el caso estudiado durante el proyecto, el marcado se realiza de forma automática mediante un programa comercial, el extractor utilizado es el que está basado en proporciones faciales [24,25,26,27], el cual solo depende de las coordenadas del centro de los ojos.

Este segmentador, *basado en las proporciones de la cara*, extrae los rasgos faciales de interés (las cejas, los ojos, la nariz, la boca, la barbilla, la frente y las orejas) de forma exacta pero no precisa, es decir, localiza los rasgos pero no optimiza el segmento centrándolo ni alineándolo. A continuación, se explican, en la Tabla 9, las proporciones utilizadas (todas ellas dependiente de la distancia entre el centro de los ojos (IPD), considerada ideal):

Id	Línea	Proporción
-	Ancho	$IPD * 5/2$
-	Alto	$Ancho * 3/2$
-	Centro (de la cara)	$(Coordenada\_x\_pupila\_derecha + Coordenada\_x\_pupila\_izquierda)/2$
S	Izquierda	$Centro + Ancho/2$
Q	Derecha	$Centro - Ancho/2$
R	Arriba	$Coordenada\_y\_pupila\_derecha - Alto/2$
T	Abajo	$Coordenada\_y\_pupila\_derecha + Alto/2$
H	Labial (línea inferior continua)	$(3 * Alto/4 + 2 * Alto/3)/1.93$
G	Línea discontinua inferior	$Alto/3$
B	Línea discontinua superior	$2 * Alto/3$
A	Línea discontinua superior	$Alto/4$
F	Línea discontinua media superior	$Alto/2$
I	Línea discontinua media inferior	$3 * Alto/4$
J	Línea discontinua inferior	$7 * Alto/8$
D	Línea ceja media (Lcm)	$(Alto/3 + Alto/2)/2$
E	Línea ceja inferior	$Lcm + 4/3 * (Lcm - abs((Lcm + Alto/2)/2))$
C	Línea ceja superior	$Lcm - 4/3 * (Lcm - abs((Lcm + Alto/2)/2))$
K	Línea derecha exterior	$Ancho/5$
L	Línea derecha medio (pupila derecha)	$Coordenada\ x\ pupila\ derecha$
M	Línea derecha interior	$2 * Ancho/5$
P	Línea izquierda exterior	$3 * Ancho/5$
O	Línea izquierda medio (pupila izquierda)	$Coordenada\ x\ pupila\ izquierda$
N	Línea izquierda interior	$4 * Ancho/5$

Tabla 9. Tabla de proporciones (el color de la letra hace referencia al color de las líneas).

La siguiente imagen (Figura 46) representa las líneas explicadas en la tabla anterior.

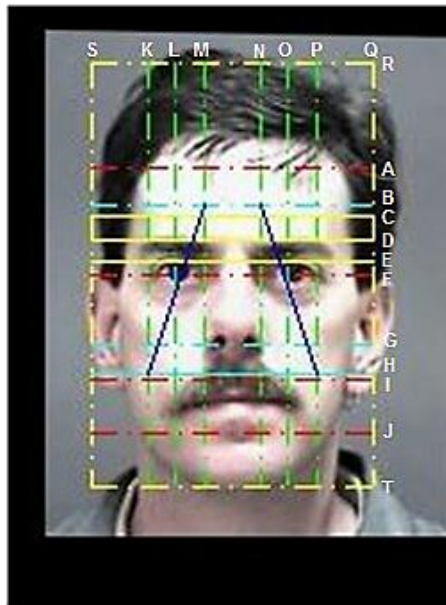


Figura 46. Proporciones faciales utilizadas.

Este tipo de extractor tiene como entrada únicamente las coordenadas de los centros de los ojos, es decir, muy útil para entornos no controlados, donde se hace complicada la localización del resto de puntos característicos de la cara.

Aplicando este segmentador sobre una imagen de la base de datos MORPH, el resultado que se genera es el siguiente (Figura 47).



Figura 47. Resultado del segmentador basado en proporciones faciales.



#### 4.6.2 *Extractor de rasgos faciales por puntos*

Adicionalmente, se introduce el desarrollo de un *extractor por puntos*, debido a su utilidad futura. El extractor por puntos, tiene como salida imágenes normalizadas (los ojos, las cejas, la nariz, la boca, la frente y la barbilla), es decir, todos los rasgos comunes de los usuarios de la base de datos (por ejemplo, todas las bocas), tienen el mismo tamaño y están centradas/alineadas en el punto medio de cada región de interés.

Antes de explicar el procedimiento, es preciso comentar que los puntos se consideran ideales y sin ningún error, ya que, o bien anteriormente, se ha pasado el detector y el corrector de puntos, o bien han sido marcados por un examinador humano manualmente.

Este segmentador tiene como desventaja que necesita como entrada, además de la imagen, todos los puntos característicos de la cara, que delimitan cada rasgo a extraer (en nuestro caso 14 puntos, incluyendo el punto que marca el final de la cabeza del usuario); sin embargo, su mayor ventaja es que la extracción del rasgo se hace de una manera más certera y precisa, y las imágenes de las regiones faciales resultantes están normalizadas.

Para realizar el extractor por puntos, primero se calcula el punto medio de los puntos característicos situados en los extremos de cada rasgo. Desde la coordenada del punto medio calculado anteriormente, se realiza un desplazamiento, tanto en sentido horizontal como en vertical, de  $\pm$  un offset (dependiente de la distancia entre ojos (IPD)), distinto para cada rasgo, pero el mismo para todos los usuarios, consiguiendo de esta manera, que todos los rasgos comunes de los usuarios (por ejemplo, todas las narices) tengan las mismas dimensiones y estén centradas/alineadas en el punto medio de cada región de interés. Finalmente, se extraen, de la imagen original, cada uno de los rasgos característicos de la cara de forma individualizada.

En la Tabla 10, se muestran los offsets (iguales para todos los usuarios para rasgos comunes) obtenidos, en píxeles, para las imágenes normalizadas.

Rasgos facial	Offset eje horizontal	Offset eje vertical
Frente	IPD (x2)	2* IPD /3 (x2)
Ceja derecha	IPD /2 (x2)	IPD /3 (x2)
Ceja izquierda	IPD /2 (x2)	IPD /3 (x2)
Ojo derecho	IPD /3 (x2)	IPD /3 (x2)
Ojo izquierdo	IPD /3 (x2)	IPD /3 (x2)
Nariz	2* IPD /4 (x2)	IPD + IPD /3
Boca	2* IPD /3 (x2)	IPD /3 (x2)
Barbilla	1.2* IPD (x2)	IPD /2 (x2)

Tabla 10. Offsets para las imágenes normalizadas, en píxeles.

En la Figura 48, se muestra el resultado del extractor es el siguiente para un usuario de la base de datos MORPH.

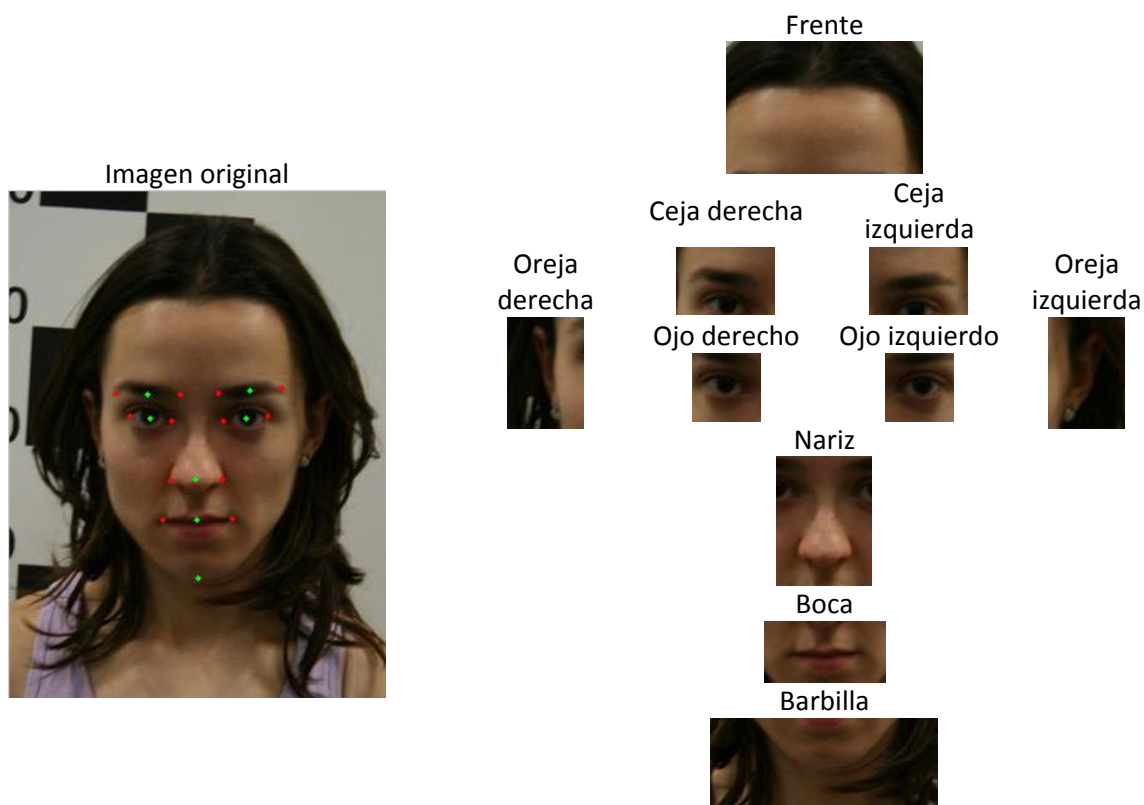


Figura 48. Resultado segmentador por puntos (ATVS).

La *principal diferencia* entre ambos extractores de rasgos, es que el extractor por proporciones se puede utilizar en cualquier entorno, controlado o no controlado, a partir de las pupilas únicamente, y el de por puntos solamente sirve para entornos controlados y necesita todos los puntos característicos marcados, teniendo mayor carga computacional.

A continuación, en la Figura 49, se realiza una comparación entre los dos segmentadores explicados en las dos secciones anteriores (el usuario utilizado para este ejemplo pertenece a la base de datos de MORPH).



Figura 49. Comparación resultado de ambos segmentadores.

Se observa que el segmentador de la izquierda (por puntos), tiene una mayor precisión y las imágenes salen centradas y alineadas, mientras que con el otro extractor no ocurre lo mismo.

### 4.6.3 Máscaras de ruido

Para mejorar los resultados del marcador/corrector desarrollado en el proyecto, se aplican *máscaras*, dependiendo del rasgo a examinar, se utilizan una u otra máscaras, afinando de esta manera el marcado y eliminando el posible ruido. Estas máscaras tapan las regiones de las imágenes que no nos interesan, dejando solo el rasgo en sí; de esta manera, se disminuye el posible error en el marcado.

La máscara elíptica (Figura 50) se ha utilizado para los ojos, las cejas, la boca y la cara (aunque de distinta forma). Se pueden modificar los valores de entrada de la función que realiza la máscara para hacer más o menos ancha la elipse o centrarla en un punto en concreto.



Figura 50. Máscaras elípticas.

La máscara triangular se aplica a la nariz, tapando las mejillas, parte de los ojos y parte de las cejas mostradas en el segmento extraído. Se puede modificar el valor de entrada de la función que realiza la máscara para hacer más o menos ancha la parte superior de la máscara, para que se pueda aplicar a todo tipo de narices (delgadas, anchas...)



Figura 51. Máscaras triangulares.

Como en la práctica es muy difícil diferenciar el tipo de narices sin los puntos característicos, la máscara triangular utilizada es la representada bajo el título de nariz media (Figura 51).

La última máscara desarrollada para el proyecto, es la máscara de dos pendientes (Figura 52). Ésta se aplica a las mitades de la cara (parte izquierda y derecha de la cara), tapando el fondo de la imagen (que está detrás de la cabeza), parte del pelo y el cuello, mostrados en el segmento extraído. Esto se realiza mediante dos rectas superpuestas con distintas pendiente (una vertical y otra horizontal).



Se pueden modificar los valores de entrada de la función que realiza la máscara para modificar la pendiente y el ancho de la parte superior e inferior de la de la recta vertical y vertical.



Figura 52. Máscaras dos pendientes.

#### 4.6.4 Realce, ecualización y umbralización

Una vez localizado el segmento del rasgo, se aplican distintas operaciones explicadas a continuación:

- Realce (Figura 53)

La función de matlab utilizada para realizar un realce en la imagen es  $J = \text{imadjust}(I, [LOW\_IN; HIGH\_IN], [LOW\_OUT; HIGH\_OUT])$ , que asigna los valores de la imagen de intensidad  $I$  a nuevos valores en  $J$  de manera que los valores entre  $LOW\_IN$  y  $HIGH\_IN$  corresponden a los valores entre  $LOW\_OUT$  y  $HIGH\_OUT$ . Los valores por debajo de  $LOW\_IN$  y por encima de  $HIGH\_IN$  se recortan, es decir, los valores entre  $LOW\_IN$  y  $LOW\_OUT$ , y los valores entre  $HIGH\_IN$  y  $HIGH\_OUT$ . Se puede utilizar una matriz vacía ( $[]$ ) para  $[LOW\_IN; HIGH\_IN]$  o para  $[LOW\_OUT; HIGH\_OUT]$  con el fin de especificar el valor predeterminado de  $[0 \ 1]$ . Si se omite el argumento,  $[LOW\_OUT; HIGH\_OUT]$  defecto  $[0 \ 1]$ .



Figura 53. Realce.

En el ejemplo anterior:

- $LOW\_IN=0.3$ ,  $HIGH\_IN=0.7$
- $LOW\_OUT=0$ ,  $HIGH\_OUT=1$

Con estos valores se consiguen resaltar los rasgos faciales con respecto a la cara.

- Ecualización (Figura 54)

La ecualización mejora el contraste de imágenes mediante la transformación de los valores de una imagen de intensidad (imagen cuyos valores de píxeles corresponden a una escala de grises), o los valores en el mapa de colores de una imagen indexada, de manera que el histograma de la imagen de salida coincide aproximadamente con un histograma especificado. La función utilizada en matlab para la ecualización es  $J = \text{histeq}(I)$ . Esta función devuelve la transformación de escala de grises que asigna niveles de gris de la imagen de intensidad  $I$  a niveles de gris de  $J$ .



Figura 54. Ecualización.

- Umbralización (Figura 55)

La umbralización es un caso particular del realce, con la diferencia de que en este caso, LOW\_IN y HIGH\_IN toman el mismo valor; dejando la imagen bimodal (blanco y negro).



Figura 55. Umbralización.

Se consigue con esto, optimizar para la última etapa de localización.

#### 4.6.5 Barrido, localización y marcado de puntos

Una vez extraído los segmentos de interés de la cara, mediante el segmentador basado en proporciones faciales (debido a que se utiliza una base de datos no ideal), y haber aplicado la máscara correspondiente sobre estos, se procede a la búsqueda de los nuevos puntos característicos de dicho rasgo.

Para la localización de los puntos, se utiliza, como base, el mismo algoritmo, pero dependiendo del rasgo a estudiar, se añaden operaciones aparte para optimizar el resultado.

El método básico y principal utilizado para la localización de puntos característicos de la cara, consiste en realizar un *barrido vertical* por toda la imagen (tal y como se muestra en la Figura 56) hasta encontrar el punto de interés, ya que, estos puntos son los extremos del rasgo (si se buscase un punto más bajo y más alto, por ejemplo para ver la altura del rasgo, el barrido a utilizar sería el horizontal).

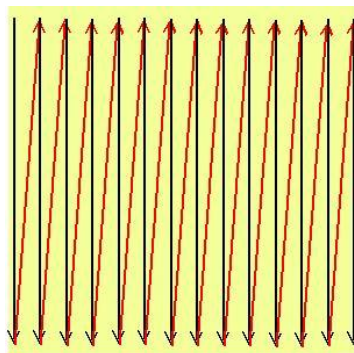


Figura 56. Barrido vertical

Como punto de partida, este algoritmo utiliza una imagen que ha sido tratada anteriormente, es decir, no se utiliza la foto original, sino que se ha extraído el rasgo a estudiar, se ha realizado un realce de contraste (opcional, ya que no siempre es necesario), se han detectado los bordes y, finalmente, se le ha aplicado la máscara correspondiente. Posteriormente, el barrido comienza desde el primer píxel de la imagen obtenida, situado en la esquina superior izquierda de ésta, y se comprueba si hay píxeles blancos en todas las filas columna a columna. Después, se analiza si en la columna a estudiar, hay alguna fila con su píxel en blanco; en caso negativo, se continúa con el barrido hasta que se encuentre la primera columna con píxeles blancos, en caso afirmativo, el algoritmo de barrido ya ha encontrado un extremo del rasgo a estudiar.

El primer extremo del rasgo a estudiar es el punto medio entre todos los píxeles blancos encontrados en la columna. Este es el procedimiento a seguir, para evitar que una impureza (en este caso, barba, bigote, arrugas...), desplace en exceso el punto a obtener. Para obtener el otro extremo del rasgo, se voltea la imagen, obtenida al principio, de izquierda a derecha (función matlab `fliplr`), y se repiten los mismos pasos explicados. El resultado de aplicar el algoritmo explicado sobre el ojo derecho de un usuario de la base de datos BioID, se muestra en la Figura 57.



Figura 57. Resultado algoritmo barrido vertical sobre ojo.

Este barrido vertical, se puede aplicar ya que, debido a la segmentación, mediante el extractor por proporciones faciales pues estamos en un entorno no controlado, y a las máscaras aplicadas a cada uno de los rasgos, la parte de la cara a estudiar está muy restringida y optimizada al rasgo, eliminando todo tipo de ruido.

Para mejorar esta localización, además del barrido vertical, se realizan diversas operaciones adicionales, como:

- Realzar el contraste: se aplica cuando la iluminación de la imagen no es la adecuada, ya sea porque es excesiva o escasa; también se aplica en pieles oscuras, como por ejemplo, las personas africanas.
- Ecualización del histograma de la imagen: consiste en realizar una transformación que pretende obtener para una imagen un histograma con una distribución uniforme. Es decir, que exista el mismo número de píxeles para cada nivel de gris del histograma de una imagen monocroma.
- Eliminar impurezas, es decir, un conjuntos de píxeles blancos (en una imagen binarizada) de tamaño 1 o 2.

Un caso especial, es la boca, ya que la obtención de los puntos característicos de este rasgo se hace muy complicada, por ejemplo, en gente con barba, bigote, obesa, con piel ajada, con el labio hendido...

Por ello, además del barrido vertical, se aplica a este rasgo otro método, representado el resultado en la Figura 58, que consiste en, una vez obtenido el segmento de la cara a estudiar, es decir, la boca, ecualizar el histograma de dicha imagen, para, posteriormente, binarizarla con un umbral muy bajo (0.2). Después, se dilata la imagen resultado una sola vez, aumentando la cantidad de píxeles en blanco (con valor 1) de dicha imagen, se eliminan los píxeles interiores, estableciéndose un píxel a 0 (negro) si la totalidad de sus cuatro píxeles vecinos están conectados a 1 (blanco), dejando sólo los píxeles del contorno de la imagen. En este paso se pueden obtener más de un contorno (o líneas contorno), pues si la imagen es muy ruidosa y está llena de impurezas, se hace impracticable la obtención del contorno. Para solucionar el problema de varios contornos, se etiquetan las líneas contorno obtenidas, cada una con un número. Este etiquetado se realiza de la siguiente forma:

- Se establece un píxel a 1 si la totalidad de sus ocho píxeles vecinos están conectados y son de color blanco.
- Si no hay más píxeles (además de los etiquetados a 1) en toda la imagen, es que solo hay un contorno. En caso contrario, el nuevo píxel se etiqueta como 2, repitiéndose el paso anterior hasta que no haya píxeles en la imagen sin etiquetar.
- Los píxeles de color negro son etiquetados con el número 0.

Como el vecindario es muy grande (ocho píxeles), normalmente va a haber un contorno grande, y tal vez, varios de menor tamaño. Este contorno único, o en su defecto, el mayor, es eliminado de la imagen (poniendo sus píxeles a 0). Finalmente, se le aplica una máscara elíptica. A esta imagen resultado, es a la que se le aplica el barrido vertical.

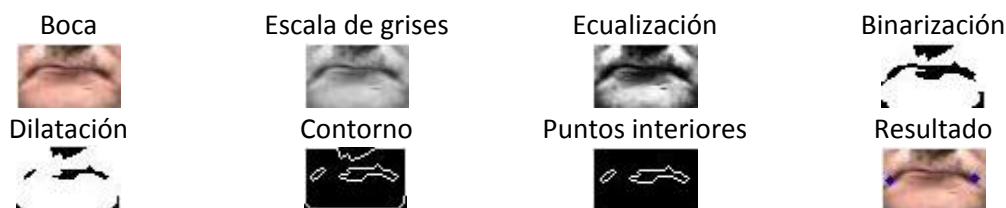


Figura 58. Resultado algoritmo barrido vertical + método especial para la boca.

La imagen denominada “Puntos interiores”, en la Figura 58, es la que utiliza el algoritmo barrido vertical.

Si no se combinasen ambos métodos, el resultado sería bastante peor, como se puede ver en la Figura 59.



Figura 59. Comparativa puntos característicos de la boca

Otro rasgo facial que presenta bastantes complicaciones es la barbilla (o mentón), por lo que también se ha optado por combinar varios métodos para la obtención de su punto característico. El punto de la barbilla, a priori, es algo más complicado que el resto de puntos debido a que no se aprecia un cambio brusco de luminancia y crominancia (debajo de la barbilla está el cuello, que suele ser del mismo color que la barbilla), en el caso de que se intente localizar dicho punto mediante el color; y puede encontrarse ocluida por la barba.

El método utilizado para la obtención del punto característico de la barbilla (explicado a continuación), es tolerante a las posibles oclusiones de la barba.

Por ello, además del método que consiste en realizar un barrido vertical, se aplica a este rasgo el siguiente método matemático, el cual parte de la imagen normalizada y rotada. Al principio, se predice que el punto de la barbilla se va a encontrar a  $1.69 \times \text{distancia\_entre\_ojos}$  píxeles hacia abajo desde las coordenadas “y” de las pupilas (son iguales la de ambas pupilas debido a la rotación). Posteriormente, se realiza un tratamiento de la imagen; es decir, se binariza y se le aplica una máscara elíptica. En la imagen obtenida, se busca, en un vecindario de  $17 \times 9$  píxeles (alto y ancho respectivamente) y centrado en el punto predicho, los píxeles de color blanco, de la siguiente manera:

- Se guardan los valores de todos los píxeles, que se encuentran dentro del vecindario, en una matriz.
- Para localizar exactamente el punto buscado, se distinguen tres casos:
  - Si la matriz no contiene ningún píxel cuyo valor es 1 (blanco), el punto buscado se mantiene en el punto predicho anteriormente.
  - Si la matriz solo contiene píxeles con valor 1, es decir, blancos, el punto final se convierte en el punto predicho más 8 píxeles.
  - Si la matriz tiene píxeles de ambos valores, 0 y 1, la coordenada y deseada es la media entre todas las coordenadas “y” finales de cada columna. Estos puntos finales de cada columna, se encuentran mediante el barrido vertical.

La coordenada “x” del punto característico de la barbilla, se considera que siempre está colocado en el punto medio que forman las pupilas de los ojos. A modo comparativo, en la Figura 60 se observa que el punto obtenido final (color amarillo) está mejor localizado que el obtenido de forma matemática (color rojo).



Figura 60. Comparación punto característico de la barbilla.

Finalmente se aplica una desnormalización para que los puntos vuelvan a corresponderse con la imagen original.

Este paso consiste en hacer la rotación inversa realizada y escalar tanto las imágenes como los puntos, para obtener una base de datos final idéntica a la original, salvo los puntos corregidos o los nuevos puntos obtenidos, como el punto que marca el final de la cabeza. La Figura 61 muestra un ejemplo de imagen final perteneciente a un usuario africano de la base de datos MORPH con todos los puntos característicos, de interés, de la cara.



Figura 61. Puntos característicos sobre un usuario de la base de datos MORPH.

Finalmente, después de aplicar el algoritmo completo, y la desnormalización, se obtiene la base de datos final (MORPH), mostrada a continuación en la Figura 62, donde hay una mejoría notable.

Los puntos de color rojo, son las coordenadas de los puntos característicos de la cara dados por el programa comercial Luxand, mientras que los marcados en color azul, son los corregidos finales.



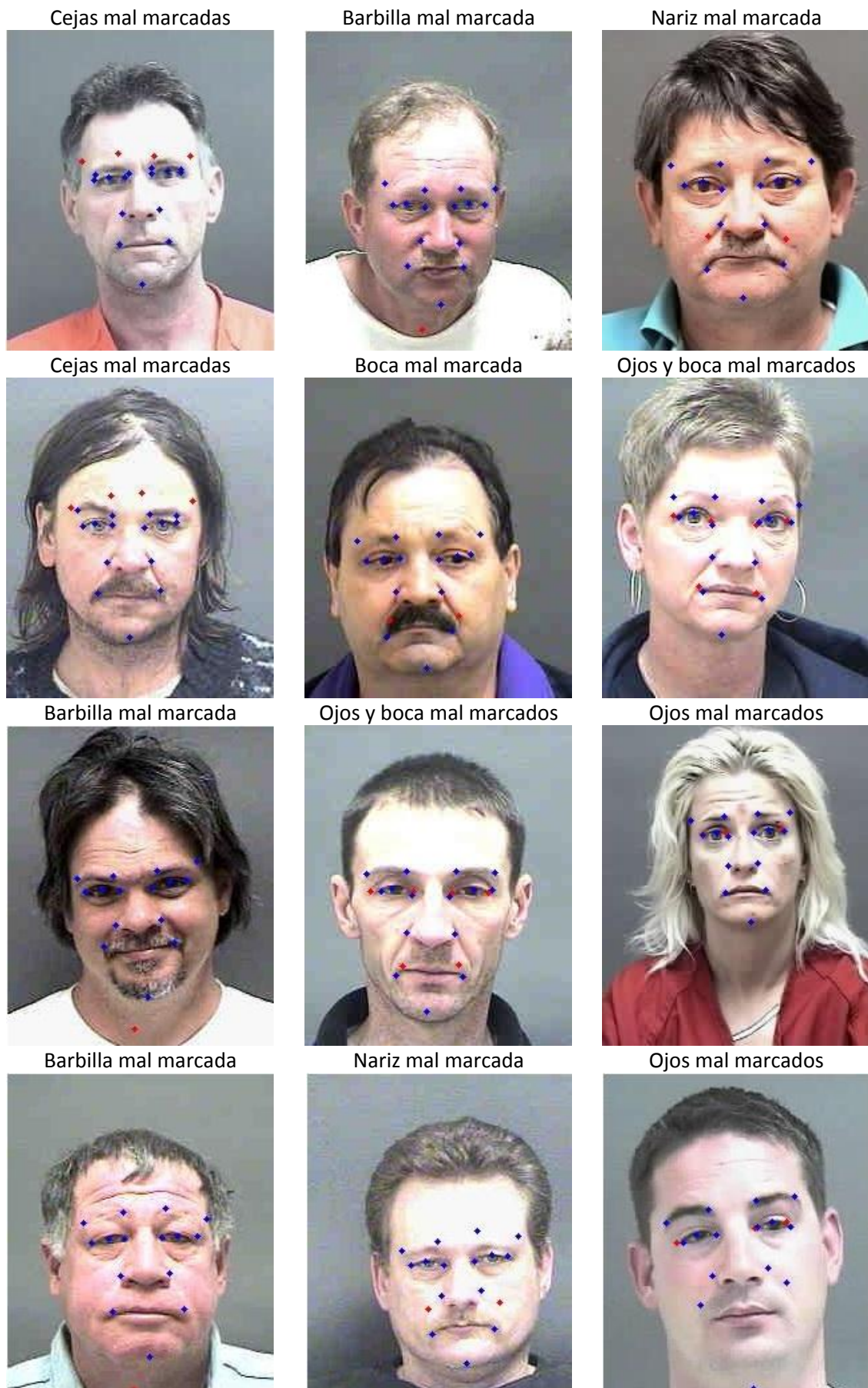


Figura 62. Base de datos MORPH corregida.

## 5. Experimentos y resultados

### 5.1 Evaluación del rendimiento de sistemas biométricos

Para evaluar las prestaciones o rendimiento de un sistema biométrico, se deben analizar y valorar los siguientes parámetros estándares:

- *FAR* (*False Reject Rate* o falsos positivos): porcentaje de personas no autorizadas aceptadas por el sistema.
- *FRR* (*False Reject Rate* o falsos negativos): porcentaje de personas autorizadas no aceptadas por el sistema.
- *EER* (*Equal Error Rate*): los indicadores FAR y FRR son parámetros inversamente proporcionales, y variarán en función de las condiciones prefijadas por el programa de identificación biométrica. Si el algoritmo se va a utilizar en un sistema de seguridad, se ha de intentar que el FAR sea lo más pequeño posible, aunque esto signifique el incremento del FRR. Se debe fijar un umbral que permita igualar estos dos factores, momento en el que se considera óptimo el funcionamiento del sistema. Este umbral se denomina *Equal Error Rate* (Figura 63) y es el que determina el poder de identificación del sistema.

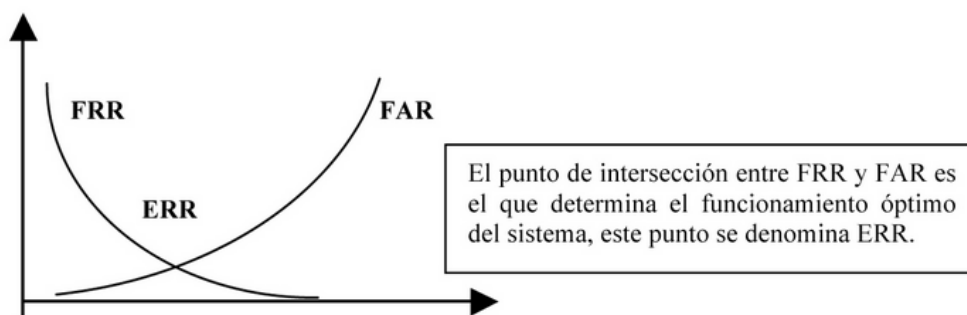


Figura 63. Equal Error Rate.

Cuanto más bajo es el EER, se considera más exacto el sistema.

En el ámbito forense, se realizan dos tipos de cotejos para comparar imágenes:

- *Morfológico*: se realiza con el fin de efectuar la comparación de los rasgos faciales y corporales de una persona con respecto a otra que aparece en imágenes de vídeo, fotografías o imágenes digitalizadas. Igualmente, señales particulares, prendas de vestir y accesorios que porte la persona en las imágenes suministradas. Una vez comparado cada uno de los rasgos morfológicos faciales, se establece el grado de semejanza de acuerdo al porcentaje resultado (Tabla 11).



Porcentaje	Grado de semejanza
1% – 19%	No hay semejanza
20% - 39%	Semejanza ligera
40% - 59%	Semejanza aproximada o aceptable
60% - 79%	Semejanza cercana
80% - 100%	Gran semejanza

Tabla 11. Grado de semejanza.

- *Antropométrico*: sirve para identificar a las personas mediante la obtención y clasificación de ciertas medidas del cuerpo humano. En España se efectuó por primera vez en el gobierno civil de Barcelona el 1 de agosto de 1895.

Este último, es el utilizado por la Dirección General de la Guardia Civil (DGGC) para la comparación de imágenes de buena calidad en la reseña fotográfica y en la identificación de DNI.

## 5.2 Análisis de características / puntos característicos

Para la comparación de imágenes, se han utilizado distintos vectores de características, tanto con el uso de las SVMs como de las distancias euclídeas:

- DGGC

Uno de los vectores de características utilizado en el proyecto, es el que emplea la DGGC para el cotejo antropométrico, en el ámbito forense; explicado a continuación.

*Puntos a tener en cuenta:*

- **Exocantio**: punto más externo de la comisura externa de la hendidura parpebral (sinónimo: ectocantio).
- **Endocantio**: punto más externo del canto interno de la hendidura parpebral (sinónimo: entocantio).
- **Alar**: punto más alejado del plano medio sagital situado sobre la aleta nasal.
- **Subnasal**: punto medio de la unión del tabique nasal.
- **Chelio**: punto más lateral de la comisura labial.
- **Estomio**: punto de encuentro de la hendidura bucal con el plano medio-sagital.
- **Gnatio**: punto más bajo del mentón en la línea sagital.

*Otros puntos de interés:*

- **Centros pupilares**, derecho e izquierdo.

La elección de los puntos más idóneos se basa fundamentalmente en la evitación de subjetividades derivadas de su variable colocación, los mencionados son fácilmente identificables para cualquier observador aún sin contar con formación en morfología facial.

Abreviaturas (Tabla 12):

Puntos	Abreviaturas
Exocantio ojo derecho	EXD
Centro pupilar derecho	CPD
Endocantio ojo derecho	END
Endocantio ojo izquierdo	ENI
Centro pupilar izquierdo	CPI
Exocantio ojo izquierdo	EXI
Alar derecho	AD
Alar izquierdo	AI
Subnasal	SN
Chelio derecho	CHD
Estomio	ES
Chelio izquierdo	CHI
Gnatio	GN

Tabla 12. Puntos característicos utilizados por la DGGC.

En la Figura 64, se muestran los puntos (de color azul), las líneas verticales y horizontales pasantes por los planos imaginarios citados anteriormente:

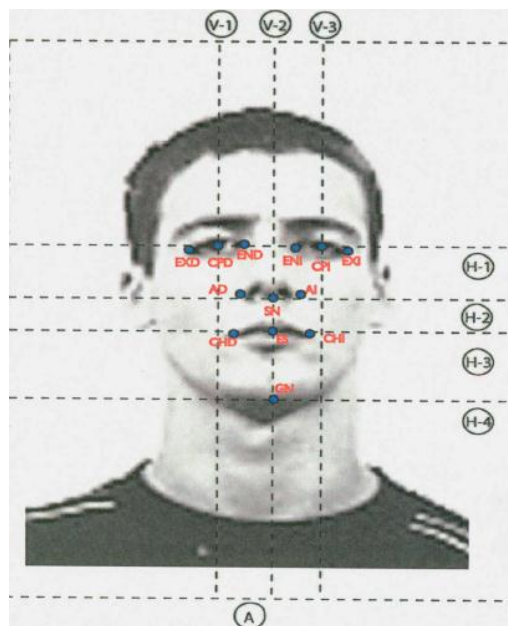


Figura 64. Puntos característicos utilizados por la DGGC.

Una vez delimitados los puntos de interés, se calculan las medidas que se van a realizar y que se ven reflejadas en la fotografía inferior (Figura 65).

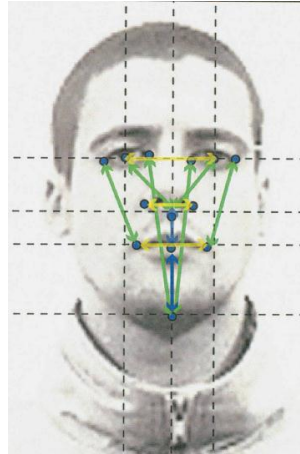


Figura 65. Mediciones a tomar.

Las medidas a tomar en sentido horizontal (amarillo) corresponden a:

- CPD-CPI
- AD-AI
- CHD-CHI

En sentido vertical (azul) corresponden a:

- SN-ES
- ES-GN

En diagonal (verde) corresponden a:

- CPD-SN
- EXD-CHD
- EXI-CHI
- CPI-SN
- END-GN
- ENI-GN

Con las distancias citadas anteriormente (11), se crea el vector de características.

- Puntos utilizados en el proyecto.

Otro de los vectores de características utilizado en el proyecto, es el que se crea con las distancias entre los puntos base y los del proyecto.

*Puntos característicos faciales de interés:*

- CDE: Punto más externo de la ceja derecha
- CDI: Punto más interno de la ceja derecha
- CIE: Punto más externo de la ceja izquierda
- CII: Punto más interno de la ceja izquierda
- ODE: Punto más externo del ojo derecho
- ODI: Punto más interno del ojo derecho
- OIE: Punto más externo del ojo izquierdo

- OII: Punto más interno del ojo izquierdo
- NOD: Punto más externo del orificio nasal derecho
- NOI: Punto más interno del orificio nasal izquierdo
- BLD: Punto más a la derecha de la boca
- BLI: Punto más a la izquierda de la boca
- M: Punto más bajo de la cara (barbilla)

Los centros pupilares son puntos dependientes:

- *Centro pupilar derecho* =  $\frac{(ODE+ODI)}{2}$
- *Centro pupilar izquierdo* =  $\frac{(OIE+OII)}{2}$

Abreviaturas (Tabla 13):

Puntos	Abreviaturas
Ceja derecha exterior	CDE
Ceja derecha interior	CDI
Ceja izquierda exterior	CIE
Ceja izquierda interior	CII
Ojo derecho exterior	ODE
Ojo derecho interior	ODI
Ojo izquierdo exterior	OIE
Ojo izquierdo interior	OII
Nariz orificio derecho	NOD
Nariz orificio izquierdo	NOI
Boca lado derecho	BLD
Boca lado izquierdo	BLI
Mentón	M

Tabla 13. Puntos de interés faciales.

En la Figura 66, se muestran los puntos (de color amarillo) citados anteriormente (izquierda) y las distancias que componen el vector de características (derecha).



Figura 66. Puntos característicos de la cara de interés en el proyecto.

Las distancias (78) con las que se crea el vector de características, son las que hay desde cada uno de los puntos, a los puntos restantes.

- Comparación

En la Tabla 14, se muestran las longitudes de los distintos vectores de características.

Vectores	Longitud/usuario
DGGC	11
Puntos proyecto	78
FaceSDK <sup>2</sup>	2145

Tabla 14. Longitud vectores de características.

Una vez obtenidos los vectores de características, se procede a realizar una comparación de dichos vectores que nos permite saber la probabilidad de coincidencia entre imágenes.

Se han utilizado tres comparadores:

- ❖ Comparador del sistema comercial FaceSDK

El comparador del sistema comercial FaceSDK se basa en la realización de un reconocimiento antropométrico (por distancias) más un análisis de textura. El sistema cuenta con el marcado de 66 puntos faciales.

- ❖ Distancias euclídeas

Debido a la colaboración realizada, durante el desarrollo del proyecto fin de carrera, con la Dirección General de la Guardia Civil (DGGC), se ha utilizado la distancia euclídea entre puntos característicos faciales como clasificador por semejanza con sus protocolos.

La distancia euclídea o euclidiana es la distancia “ordinaria” entre dos puntos de un espacio euclídeo, la cual se deduce a partir del Teorema de Pitágoras.

En general, la distancia euclidiana entre los puntos  $P = (p_1, p_2, \dots, p_n)$  y  $Q = (q_1, q_2, \dots, q_n)$ , del espacio euclídeo n-dimensional, se define como:

$$d_E(P, Q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$

- ❖ Clasificador SVM

Se realizaron experimentos utilizando las SVMs, pero se descartaron al ver que los resultados generados no eran aceptables. Esto fue debido a que las clases estaban desbalanceadas y a la falta de datos.

<sup>2</sup> Se han realizado pruebas con este vector de características, formado por las distancias entre los 66 puntos característicos faciales marcados por el sistema comercial FaceSDK, pero se han desechado debido a los malos resultados obtenidos.

Se intentó la utilización en los experimentos de este clasificador, debido a las múltiples ventajas que presenta:

- No hay problemas con mínimos locales (problema cuadrático).
- La solución óptima se puede encontrar en tiempo polinómico.
- Pocos parámetros libres:  $C$ , el tipo de kernel y los parámetros del kernel.
- Resultado estable (no depende de valores iniciales aleatorios).
- Solución dispersa, sólo se tienen en cuenta los vectores de soporte.
- Maximizar el margen permite controlar la complejidad con independencia del número de dimensiones.
- Buenas capacidades de generalización.

Con los experimentos explicados a continuación, se muestra el alcance o potencial de cada uno de los bloques desarrollados en el proyecto.

- ❖ El bloque corrector de puntos característicos de la cara, es utilizado en el experimento 0 como localizador de dichos puntos, con bastante buen resultado.
- ❖ El bloque detector es utilizado en la primera parte de los experimentos 2 y 3.
- ❖ La combinación de los dos bloques anteriores, se utiliza en la segunda parte de los experimentos 2 y 3, viéndose, como los resultados mejoran con su aplicación.
- ❖ El alcance de los comparadores se comprueba en el experimento 1, ya que se comparan imágenes marcadas de forma manual, quitando los posibles fallos de segmentación.

### **5.3 Experimento 0: Marcado automático de puntos característicos en entornos controlados**

---

Este experimento se realiza con la base de datos BioID, imágenes en escala de grises, en entornos del mundo real con variaciones de iluminación, tamaños de cara y distintos fondos.

Se va a realizar una comparación entre los 13 puntos característicos marcados por David Cristinacce y Kola Babalola (estudiantes de doctorado en el departamento de Imaging Science and Biomedical Engineering at the University of Manchester – ISBE) de forma manual, y los puntos obtenidos después de aplicar el corrector explicado en la sección 4.6 *Corrector de puntos característicos* (ya que se puede aplicar como localizador de *landmarks*), que es de forma automática, ahorrando, de esta forma, tiempo y recursos a la hora de realizar un marcado de puntos faciales.

Con esta comparación (Figura 67), se quiere observar la calidad del localizador de puntos característicos de la cara expuesto en este proyecto, justificando su utilización en el marcado de puntos en caso de error. El marcado se realiza a partir de los centros de los ojos, obtenidos después de realizar el pre-procesado de las pupilas, proporcionadas por un sistema comercial de marcado automático.

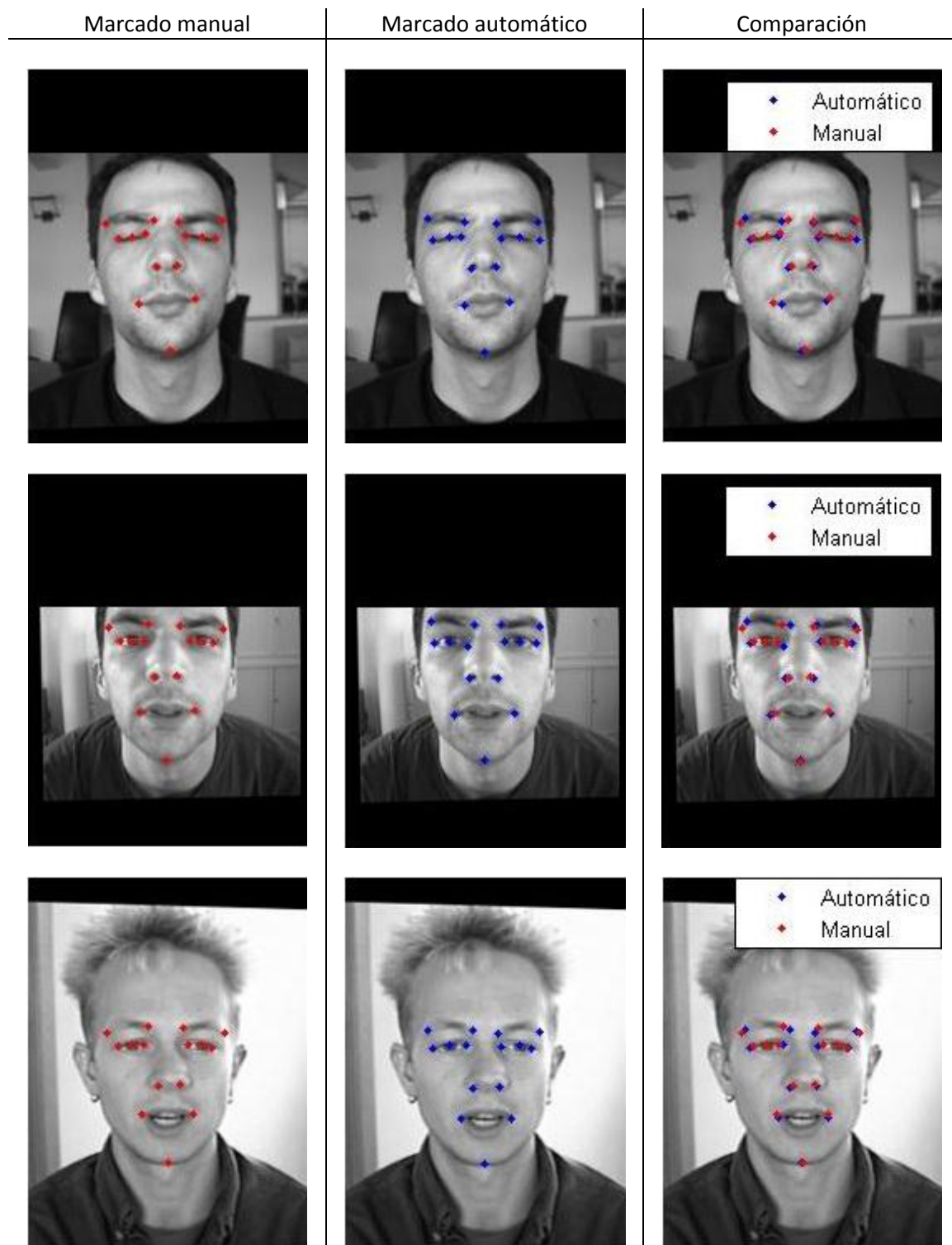


Figura 67. Resultados de marcado entre los puntos proporcionados y el sistema desarrollado en la base de datos BioID.

Visualmente, se observa que el localizador de puntos característicos tiene un margen de error pequeño; la tabla siguiente (Tabla 15), muestra con exactitud el error, en píxeles, que hay entre los puntos del sistema propio y los del sistema automático. Este error ha sido calculado como la distancia que hay entre los puntos.



Rasgo facial	Usuario 4	Usuario 1482	Usuario 1055
Ceja derecha	2.38 píxeles	4.56 píxeles	5.25 píxeles
Ceja izquierda	1.62 píxeles	2.74 píxeles	2.75 píxeles
Ojo derecho	0.69 píxeles	2.94 píxeles	1.92 píxeles
Ojo izquierdo	4.41 píxeles	2.95 píxeles	2.42 píxeles
Nariz	2.19 píxeles	2.06 píxeles	2.59 píxeles
Boca	3.69 píxeles	2.31 píxeles	4.29 píxeles
Barbilla	0.73 píxeles	0.46 píxeles	0.22 píxeles

Tabla 15. Error sistema manual vs sistema automático.

En la Figura 68, se representa el error medio de cada rasgo, para cada uno de los usuarios de la base de datos (hay que tener en cuenta, que aunque hay 1521 imágenes, realmente solo existen 23 usuarios, los cuales, tienen varias imágenes en la base de datos).

Para los rasgos faciales con dos puntos característicos (cejas, ojos, nariz y boca), el error medio se ha calculado como la media de los errores de cada uno de los puntos obtenidos, ya que si no, se da el caso de que una boca bien marcada frente a una barbilla mal marcada, tiene más error la boca, ya que se suman dos errores (uno por cada punto) mientras que la barbilla solo tiene un error.

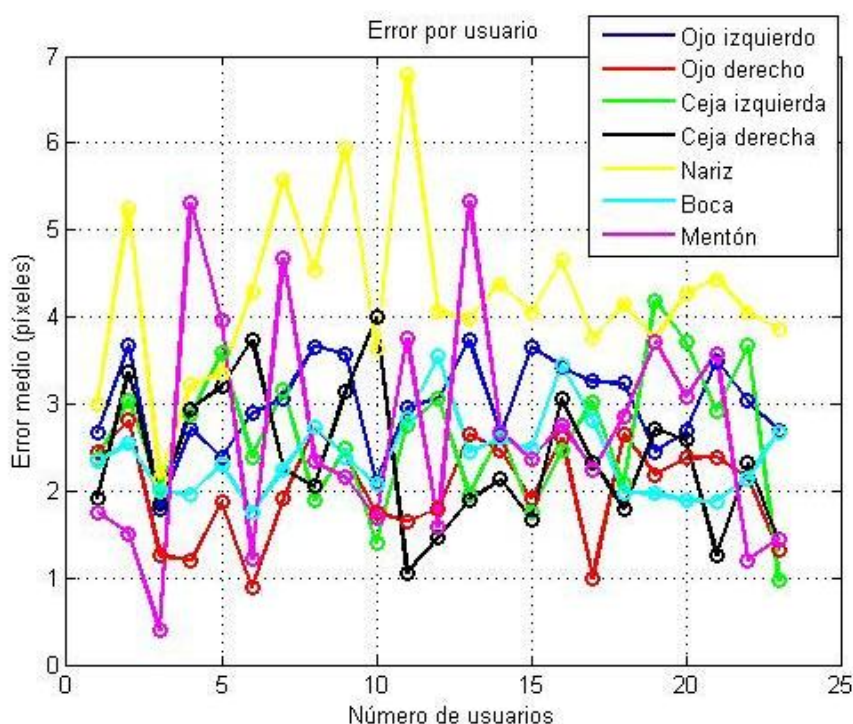


Figura 68. Error medio, de cada rasgo facial, por usuario.

Con la gráfica anterior, se pueden obtener varias conclusiones:

- El rasgo facial en el cual ha sido más complicado la localización de sus puntos característicos, es la nariz. Esto es debido a que, aunque es fácil encontrar los orificios nasales, en caras que están mirando hacia abajo, se hace imposible. En estos casos, se intenta localizar el ancho de la nariz, pero al no tener la nariz características estáticas, es decir, misma forma, misma altura, mismo ancho...se produce un error mayor.



Además, la comparación de los puntos obtenidos de la nariz se realiza con los de BioID, los cuales siempre se encuentran en los orificios nasales, por lo que el error, en los puntos localizados en el ancho de la nariz, no es exacto.

- La variabilidad del ojo derecho es menor que la del ojo izquierdo, por lo tanto, el marcado produce menor error en el primer caso.

Si se examina a cada uno de los miembros de la base de datos, se puede ver cuál es el usuario con mayor error medio. Esto se realiza en la siguiente gráfica (Figura 69):

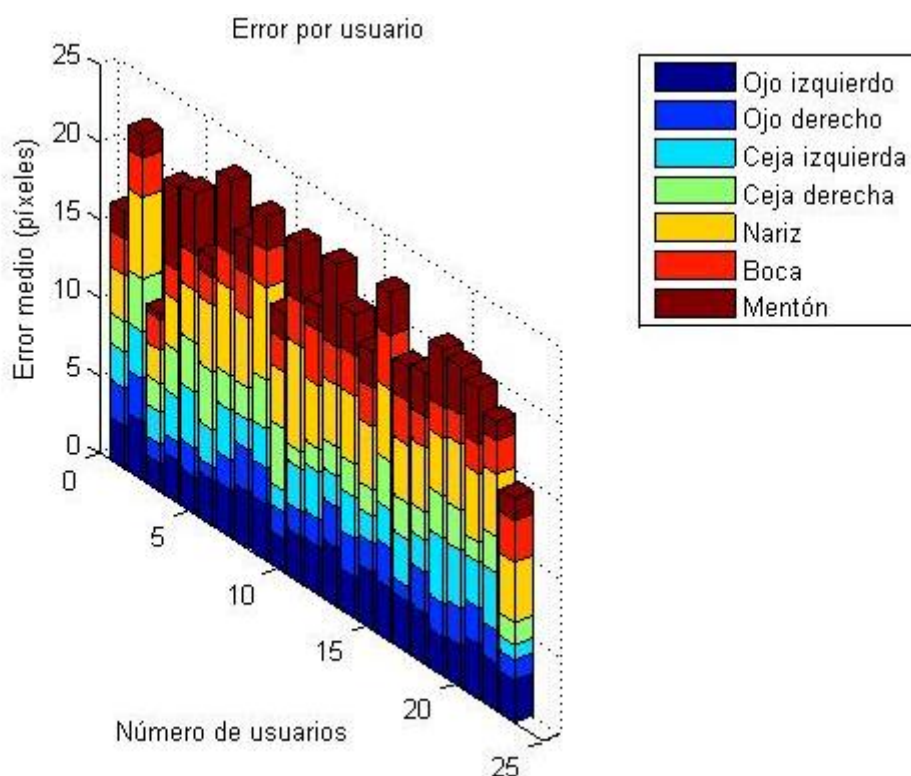


Figura 69. Error medio de cada usuario.

El usuario número tres, tiene un error medio tan pequeño debido a que solo tiene dos fotografías en la base de datos. Los usuarios con mayor error, son personas con gafas, barba, pañuelos, es decir, aquellas que presentan oclusiones o diferentes poses faciales (expresiones).

A parte de esto, algunos errores vienen dados porque los usuarios salen tapándose algún rasgo facial a examinar, lo que hace imposible la correcta localización de puntos. En la Figura 70 se observan algunos ejemplos.



Figura 70. Usuarios con rasgos faciales ocluidos.

Finalmente, se muestra en la siguiente tabla (Tabla 16) el error medio total de cada rasgo facial.

Rasgo facial	Error medio
Ojo derecho	2.06 (píxeles)
Ojo izquierdo	3.06 (píxeles)
Ceja derecha	2.29 (píxeles)
Ceja izquierda	2.54 (píxeles)
Nariz	4.36 (píxeles)
Boca	2.45 (píxeles)
Mentón	2.91 (píxeles)

Tabla 16. Error medio de cada rasgo.

Se concluye que el algoritmo presentado en el proyecto para la localización de los puntos característicos de la cara tiene como mejor resultado el ojo derecho, y como peor resultado la nariz (aunque no es exacto este error), para la base de datos dada.

#### 5.4 Experimento 1: Importancia de la segmentación

Este experimento se realiza con la base de datos MORPH (entorno semicontrolado), marcada de manera automática por el sistema comercial FaceSDK; por tanto, se considera un marcado no ideal, con error. Se ha seleccionado dicha base de datos debido a su gran cantidad de imágenes (más de 55.000 imágenes).

En la Tabla 17, se muestra, de la base de datos MORPH, el número de imágenes que tienen algún rasgo de la cara mal marcado (fases tres, cuatro y cinco), que tienen las pupilas mal posicionadas (fase dos) o en las que no se han podido encontrar los puntos característicos de la cara, incluyendo pupilas (fase uno del detector); así como el número de imágenes totales y correctas.

Hay que tener en cuenta, que solo las fases uno y dos son eliminatorias, es decir, las imágenes que no pasan alguna de estas dos fases, o ambas, se consideran incorregibles y son desechadas de la base de datos (las imágenes que fallan en las otras tres fases se consideran corregibles). Esto significa, que alguna de las imágenes puede fallar en una, dos o tres de las fases tres, cuatro y cinco.

	Fase 1 (a)	Fase 2 (b)	Fase 3 (c)	Fase 4 (d)	Fase 5 (e)	Correctas (f)	Total (g)
Africanos	29	2352	0	954	2155	37182	42589
Asiáticos	0	10	4	2	1	137	154
Europeos	4	1838	0	221	414	8098	10559
Hispanos	1	52	80	29	77	1536	1769
Otros	0	7	2	2	2	50	63
Total	34	4259	86	1208	2649	47003	55134

Tabla 17. Número de imágenes erróneas.

En la Tabla 18, se muestran los porcentajes finales de las imágenes, cuyos puntos característicos de la cara han sido corregidos (porcentaje<sub>corrección</sub>), y de las que han fallado en las fases 1 y 2, es decir, o no se ha localizado la cara en la imagen, o no se han localizado los ojos en la región de la cara adecuada (porcentaje<sub>error</sub>). Dichos porcentajes son calculados mediante las siguientes ecuaciones:

$$\text{Porcentaje}_{\text{corrección}}(\%) = \frac{g - f - (a + b)}{g}, \quad \text{Porcentaje}_{\text{error}}(\%) = \frac{a + b}{g}.$$

	Porcentaje <sub>corrección</sub> (%)	Porcentaje <sub>error</sub> (%)	Porcentaje <sub>bien</sub> (%)
<b>Africanos</b>	7.10 %	5.59 %	87.31 %
<b>Asiáticos</b>	4.54 %	6.49 %	88.97 %
<b>Europeos</b>	5.86 %	17.44 %	76.70 %
<b>Hispanos</b>	10.17 %	2.99 %	86.84 %
<b>Otros</b>	9.52 %	11.11 %	79.37 %

Tabla 18. Error de segmentación

Adicionalmente, se ha añadido una cuarta columna (Porcentaje<sub>bien</sub>) a la Tabla 18, que indica el número de imágenes que han sido marcadas por el sistema comercial de forma correcta. Este porcentaje se calcula restando al total de 100%, los dos porcentajes obtenidos anteriormente (porcentaje<sub>corrección</sub> y porcentaje<sub>error</sub>).

El objetivo de la segmentación es simplificar y/o cambiar la representación de una imagen en otra más significativa y más fácil de analizar. La segmentación se usa tanto para localizar objetos como para encontrar los límites de estos dentro de una imagen. En nuestro caso, la segmentación ha sido aplicada a la cara, para extraer cada uno de los rasgos faciales de forma independiente, permitiendo de esta manera, una corrección del marcado de puntos característicos más precisa. Como se puede deducir observando la Tabla 18, se consigue, mediante la segmentación y los dos módulos desarrollados en el proyecto (detector y corrector), un aumento medio del 7.44 % de mejora en el marcado de puntos característicos de la cara, mostrando por tanto la utilización de estos tipos de módulos desarrollados de detección/corrección que permiten afianzar los resultados.

## 5.5 Experimento 2: Reconocimiento antropométrico.

Este experimento se realiza con la base de datos *Forensic ATVS*, marcada de forma manual por una persona (no experta en temas forenses); por tanto, se considera un marcado ideal, sin ningún error.

### 5.5.1 Protocolo experimental

Esta base de datos consta de 50 usuarios, con 3 imágenes por usuario (haciendo un total de 150 imágenes). Para este experimento solo se usarán el conjunto de las fotografías que han sido capturadas a una distancia de 1 metro, con una cámara de gran calidad y una iluminación y flash adecuados (entorno controlado).

Por cada usuario, se capturaron 4 imágenes por cada sesión, para replicar los protocolos forenses. Se ha decidido utilizar solo 1 imagen de la primera sesión para train y 2 de la segunda sesión para test. Por tanto, se tiene un experimento 1:2. Tanto las imágenes de train como las de test, están realizadas bajo las mismas condiciones.

Aunque entrenar con una sola imagen es poco común (debido a que el Error Equal Rate (EER) es muy grande), se ha decidido hacerlo de esta manera, ya que, en las situaciones donde es necesario el reconocimiento facial (situaciones forenses), solo se cuenta con una imagen del usuario.

Se replica el protocolo forense en que suelen tener una imagen indubitada y varias dudosas. De manera que es un protocolo muy complejo puesto que, en reconocimiento facial, la mayoría de sistemas funcionan mejor cuanto mayor sea el número de muestras de entrenamiento.

### 5.5.2 Resultados y conclusiones

En esta segunda parte, se va a llevar a cabo el reconocimiento de patrones propiamente dicho. El reconocimiento de patrones consiste en la construcción de un sistema de clasificación a partir de un conjunto de vectores de características.

Los dos vectores de características (Figura 71) utilizados en este experimento, están formados por medidas antropométricas, es decir, las distancias que hay entre determinados puntos característicos de la cara. El primer vector, está compuesto por las 11 distancias propuestas por la DGGC, mientras que el segundo vector, está formado por las 78 distancias resultantes de los 13 puntos estudiados durante el proyecto, más 4 de las distancias utilizadas por la DGGC que no contemplábamos en nuestro vector de características, debido a que estas distancias dependen de puntos no característicos por sí mismos (dependen de otros puntos, como es el caso del punto medio de la boca, la nariz y los centros de los ojos); haciendo un total de 82 distancias.

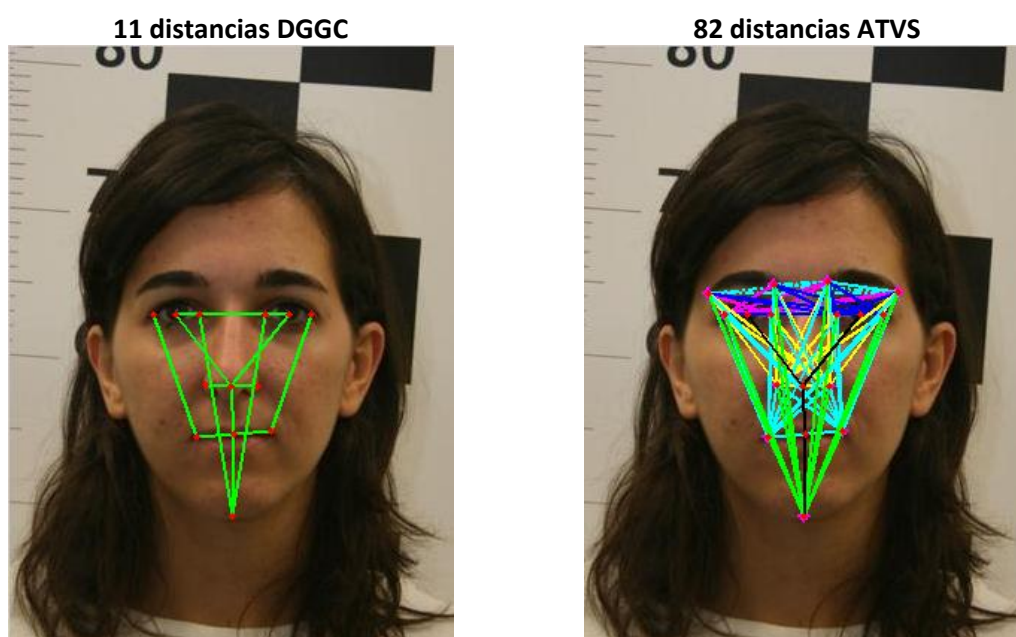


Figura 71. Vectores de características.

Después de obtener los resultados de comparación de los vectores explicados anteriormente, se ha estudiado cuáles de las características de los precedentes vectores son las más discriminantes para el reconocimiento de patrones. Para dicho estudio, se ha utilizado el algoritmo **SFFS** (Sequential Forward Floating Selection) [28] que empieza con el conjunto vacío y va incluyendo nuevas características por medio de un procedimiento secuencial hacia delante, seguido por una serie de exclusiones condicionales de la peor característica en el nuevo subconjunto de características seleccionadas. Las características permanecen o son excluidas del conjunto en función de su efectividad clasificadora. Se ha escogido este método porque es mucho más rápido que otros algoritmos existentes, como el de ramificación y acotamiento, y su eficiencia computacional es aceptable.

Estos resultados (Tabla 19) han sido obtenidos para tres tipos de marcados de puntos característicos de la cara: i) marcado manual, es decir, una persona (no experta) ha marcado, una a una, las imágenes contenidas en la base de datos; ii) marcado automático propio, es decir, mediante el bloque corrector/marcador de puntos característicos de la cara, basado en la segmentación de la cara, se han marcado las imágenes; iii) marcado automático mediante un programa comercial (después de ser corregido por el bloque corrector, desarrollado en el proyecto), es decir, el programa marca los landmarks en las imágenes, sin recurrir a la segmentación.

	Marcado manual		Marcado automático propio		Marcado automático sistema comercial		Matcher comercial
	EER original	EER SFFS	EER original	EER SFFS	EER original	EER SFFS	EER
<b>11 distancias (DGGC)</b>	18.2 %	<b>(9 dist.) 16.3 %</b>	31.9 %	<b>(4 dist.) 22.7 %</b>	40.9 %	<b>(5 dist.) 26.0 %</b>	<b>3 %</b>
<b>82 distancias</b>	13.2 %	<b>(15 dist.) 8.69 %</b>	27.9 %	<b>(15 dist.) 17.0 %</b>	28.9 %	<b>(15 dist.) 17.1 %</b>	

Tabla 19. Comparación EER

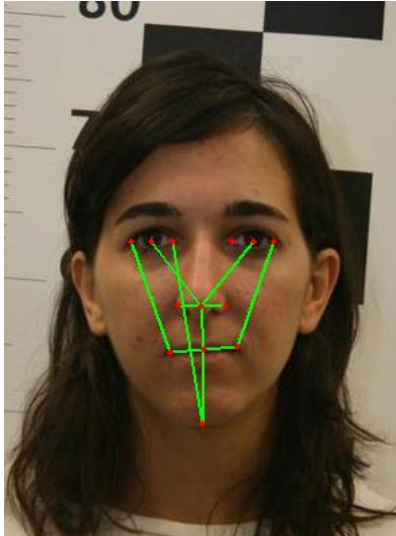
Se observa en la Tabla 19, que los resultados de EER mejoran de manera considerable después de haber hecho una selección de características mediante el algoritmo SFFS. Además, como es de esperar, el mejor resultado obtenido es cuando se ha realizado un marcado manual, sin embargo, también es el método de marcado que más tiempo requiere. Dentro de los dos marcados automáticos propuestos, el que mejor EER final tiene, aunque peor que el calculado con el marcado manual, es el desarrollado en el proyecto, el cual está basado en la segmentación de la cara en rasgos independientes (a pesar de que el marcado con el sistema comercial ha sido mejorado con la aplicación del bloque corrector). El tiempo invertido en este tipo de marcado, es bastante inferior al manual.

Otra apreciación a tener en cuenta, es que el vector de características que mejor resultado ha obtenido es el que está formado por las 82 distancias explicadas anteriormente. Esto es debido a que las distancias utilizadas, y propuesta por la DGGC, no son las más discriminantes de la cara. En la Figura 72, se muestran las distancias más discriminativas de la cara, calculadas mediante el algoritmo SFFS, para los tres métodos de marcado propuestos y ambos dos vectores de características.

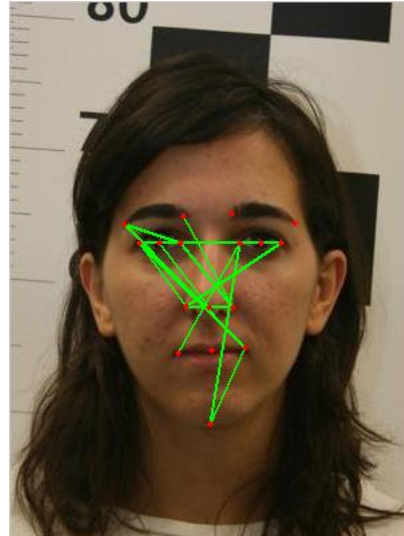


**Marcado manual**

**9 distancias discriminantes (DGGC)**

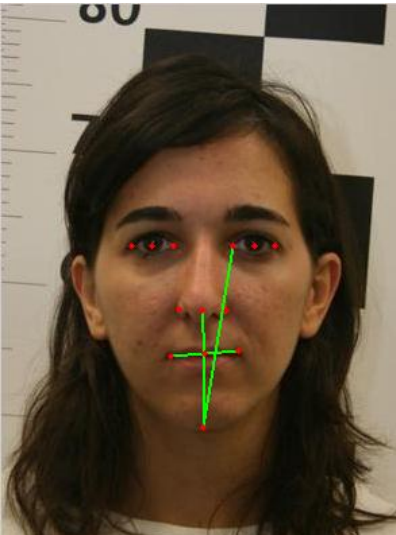


**15 distancias discriminantes (ATVS)**

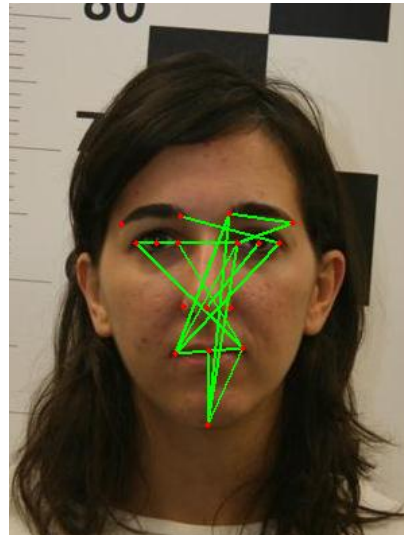


**Marcado automático propio**

**4 distancias discriminantes (DGGC)**

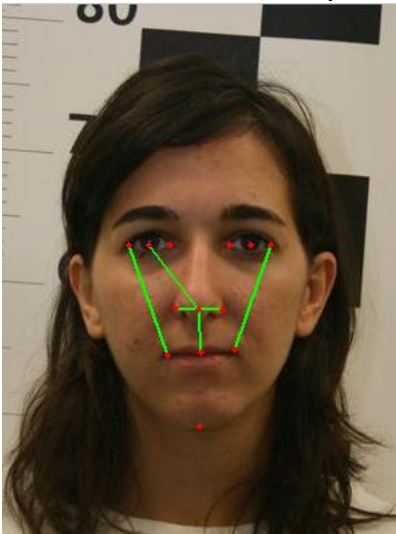


**15 distancias discriminantes (ATVS)**



**Marcado automático sistema comercial**

**5 distancias discriminantes (DGGC)**



**15 distancias discriminantes (ATVS)**

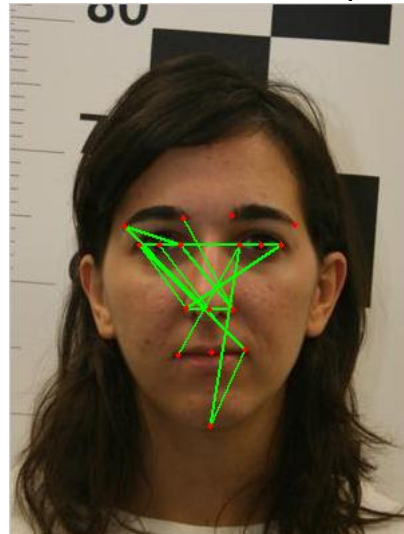


Figura 72. Distancias más discriminantes de la cara.

Adicionalmente, se ha comprobado el resultado generado por un sistema comercial del estado del arte, cuyo matcher combina distancias y textura. Como se observa en la Tabla 19 (última columna), los resultados mejoran sobre manera, pero los EERs obtenidos con el algoritmo desarrollado durante el proyecto, no están muy alejados de los mismos.

Lógicamente, al tener menor precisión en el marcado de los puntos característicos de la cara, cuando se realiza de manera automática, el número de distancias discriminantes disminuye, al haber menos ortogonalidad entre ellas.

En la Figura 73, se representa el error de segmentación medio de cada rasgo (marcado manual frente al marcado automático producido por el sistema comercial), para cada uno de los usuarios de la base de datos (hay que tener en cuenta, que aunque hay 150 imágenes, realmente solo existen 50 usuarios, los cuales, tienen 3 imágenes en la base de datos). Debido a este error de segmentación, el EER es mejor (más bajo) en marcado manual que en marcado automático.

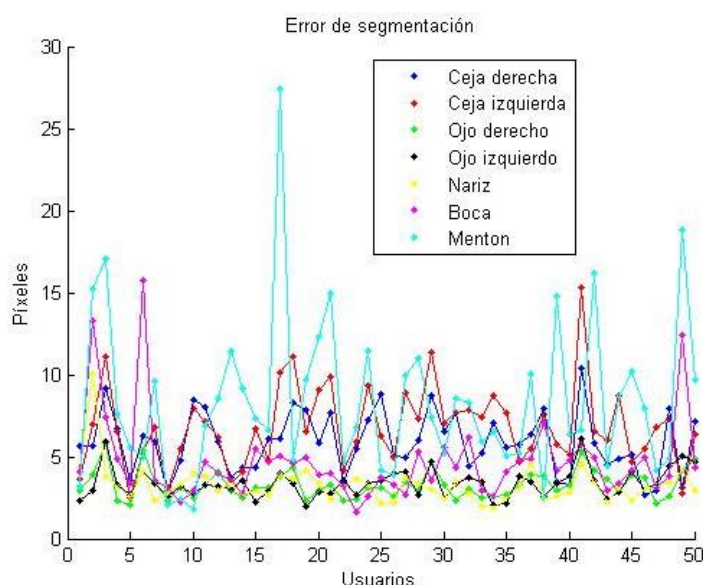


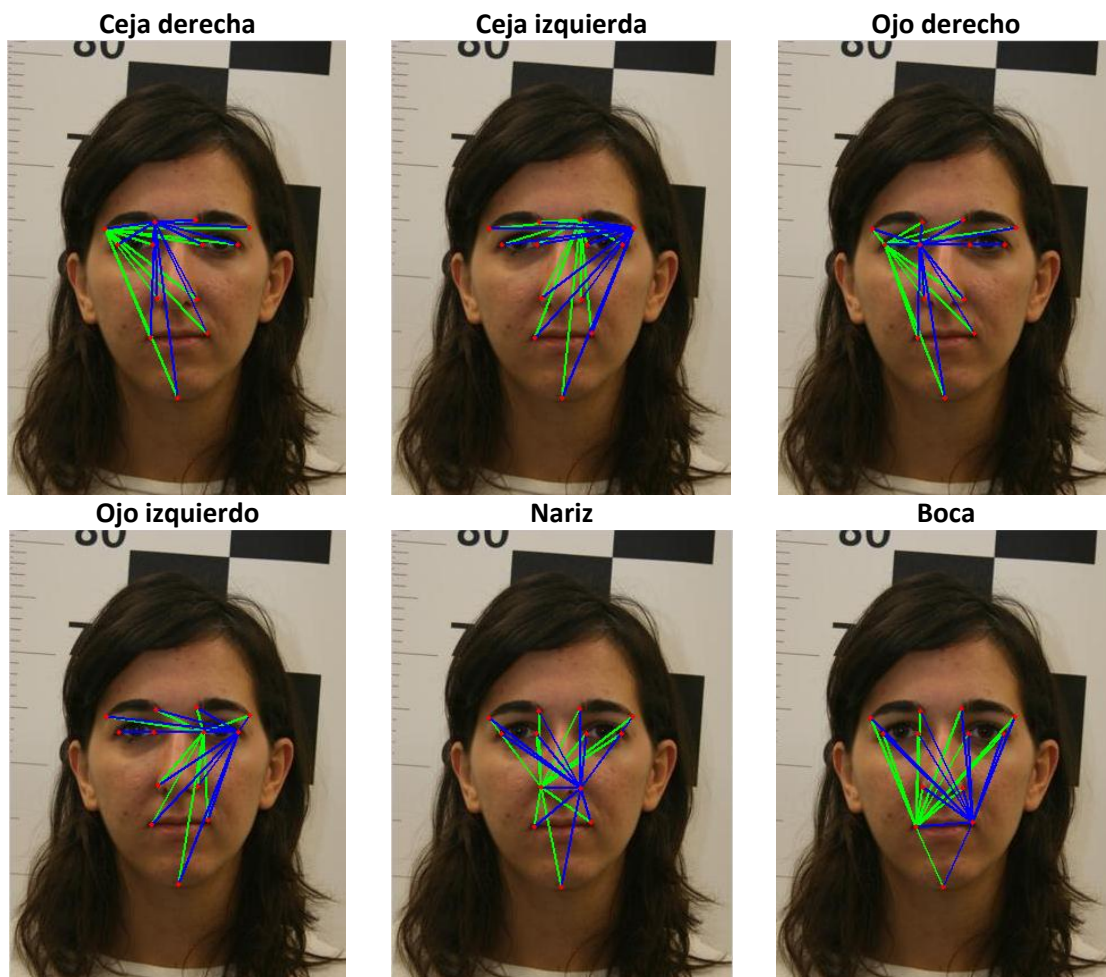
Figura 73. Error de segmentación ATVS.

Para completar el experimento, adicionalmente se ha realizado un estudio (solo con los puntos obtenidos mediante el marcado manual) comparativo entre rasgos faciales, es decir, que rasgo, solo teniendo en cuenta las distancias entre los puntos característicos del rasgo a estudiar y el resto de landmarks, es más estable/discriminante. Los resultados mostrados en la Tabla 20, son referidos al reconocimiento de los vectores de características (Figura 74) antes y después de utilizar el algoritmo de selección de características SFFS. La longitud de todos los vectores de características originales, antes de aplicar el algoritmo SFFS, es 23, exceptuando el del mentón, que es de 12 (debido a que es el único rasgo facial compuesto por un único punto característico). La última columna de la Tabla 20, hace referencia al número de distancias discriminativas, obtenidas mediante el algoritmo SFFS.

Rasgos faciales	EER original (%)	EER SFFS (%)	Números de distancias seleccionadas
Ceja derecha	13.90 %	11.26 %	12
Ceja izquierda	16.90 %	13.26 %	10
Ojo derecho	14.20 %	9.18 %	12
Ojo izquierdo	16.30 %	13.14 %	13
Nariz	18.90 %	16.00 %	12
Boca	16.90 %	16.00 %	15
Mentón	25.34 %	21.79 %	5

Tabla 20. Distancias por rasgos.

Se observa que el mentón es el rasgo menos discriminativo, a través de la comparación mediante distancias entre rasgos, de la cara; esto es debido a que las distancias del interior de la cara muestran mayor estabilidad.





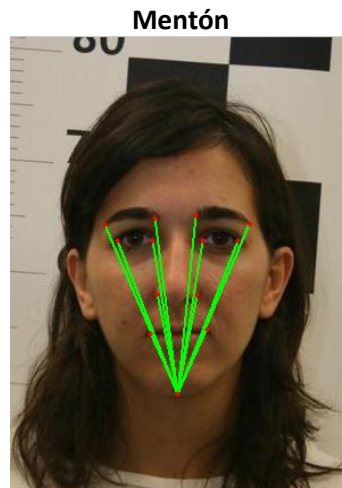
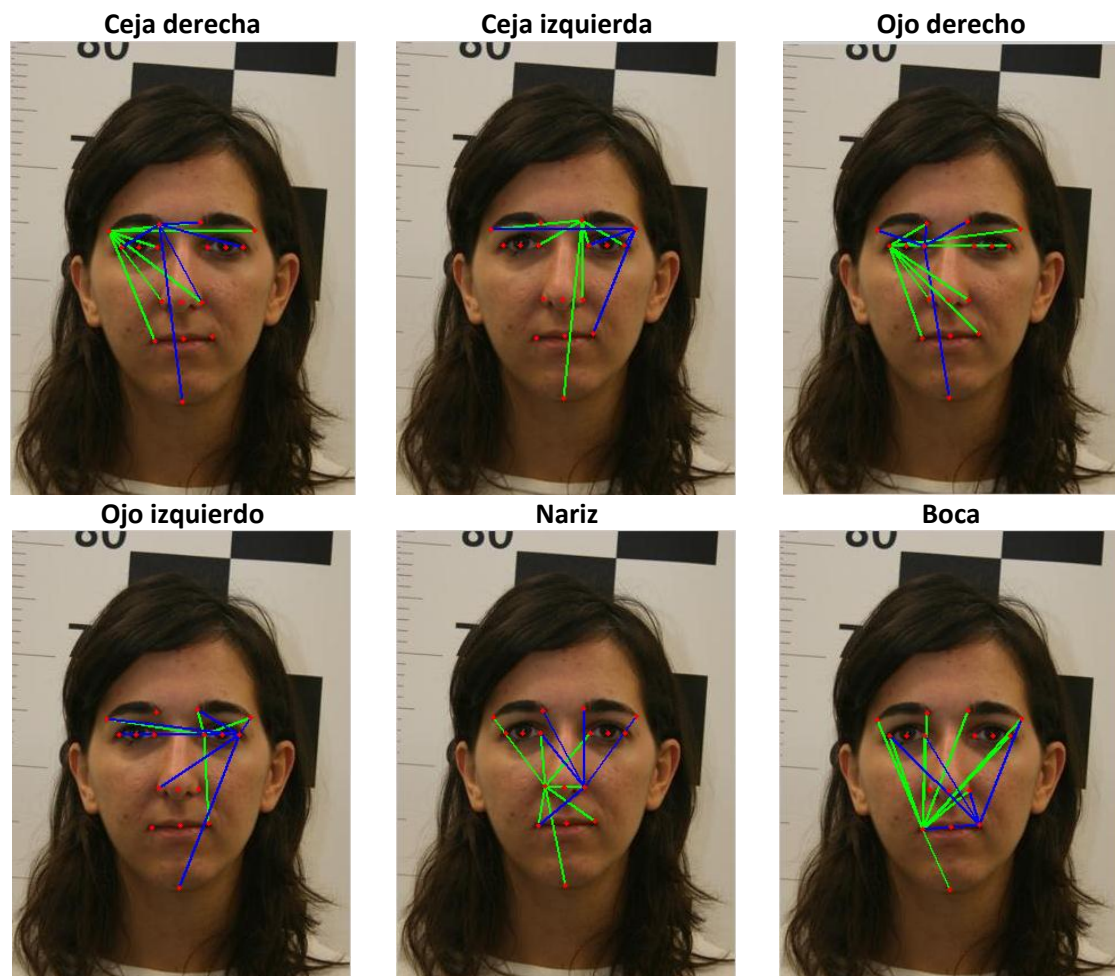


Figura 74. Vector características rasgos faciales.

En la Figura 75, se muestran las distancias más discriminativas de cada rasgo, calculadas mediante el algoritmo SFFS, para el método de marcado manual.



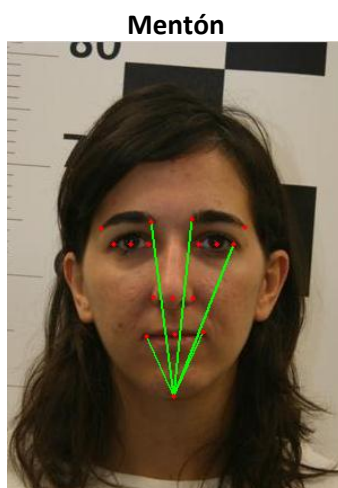


Figura 75. Distancias más discriminantes de cada rasgo.

### ***5.6 Experimento 3: Análisis del poder discriminante de cada rasgo facial para ambos tipos de extractores***

Para este experimento se han utilizado dos bases de datos: i) Scface (entorno no controlado) y ii) *Forensic ATVS* (entorno controlado), ambas marcada de manual por una personal no experta en temas forenses; por tanto, se considera un marcado ideal, sin ningún error.

En este experimento se va a comprobar el potencial de los dos extractores de rasgos faciales desarrollados en el proyecto (el basado en proporciones y el basado en puntos) en entornos tanto controlados (*Forensic ATVS*) como en entornos no controlados (*Scface*), y así poder analizar el poder discriminante de cada rasgo facial [29].

Es importante mencionar, que para los entornos no controlados, como en el caso de la base de datos *Scface*, se ha desechado aplicar el reconocimiento antropométrico debido a que, como se muestra en el experimento anterior, al aumentar el error en el marcado de puntos característicos, aumenta el error del sistema. Tanto es así, que se realizaron experimentos y este efecto se corroboró obteniendo EERs del 40%-48%. Adicionalmente, también se probó el reconocimiento antropométrico, pero en vez de utilizar las distancias euclídeas como clasificador, se utilizó el sistema de SVMs, pero debido a la falta de datos de entrenamiento, los resultados no conseguían mejorar.

#### ***5.6.1 Protocolo experimental***

La base de datos **Scface** (entorno no controlado) está formada por 130 usuarios. Esta base de datos está compuesta por imágenes frontales con una calidad buena (imágenes utilizadas para entrenamiento), y por imágenes de CCTV de baja calidad a 3 distancias posibles (4.20 metros, 2.60 metros y 1.20 metros). Para cada una de las distancias, se utilizan 5 cámaras de calidad variables. Por tanto, para el testeo se coge una imagen por usuario, distancia y cámara, quedando así, 5 imágenes por usuario. Por tanto, en este caso se tiene un experimento 1:5.

En este experimento, se genera un vector de entrenamiento y tres vectores de test, un vector por distancia (analizando adicionalmente, el efecto de la distancia). Tanto las imágenes de train como las de test, están realizadas bajo las mismas condiciones.

La base de datos **Forensic ATVS** (entorno controlado) consta de 50 usuarios, con 3 imágenes por usuario (haciendo un total de 150 imágenes). Las fotografías han sido obtenidas a una distancia de 1 metro, con una cámara de gran calidad y una iluminación y flash adecuados (entorno controlado).

Por cada usuario, se separan las fotografías en train y test. La primera imagen de cada usuario se escoge para el entrenamiento (train), dejando las dos imágenes restantes para la realización del test. Por tanto, se tiene un experimento 1:2. Tanto las imágenes de train como las de test, están realizadas bajo las mismas condiciones.

Aunque entrenar con una sola imagen es poco común (debido a que el Error Equal Rate (EER) es muy grande), se ha decidido hacerlo de esta manera, ya que, el proyecto está dirigido a un ámbito forense, donde solo se cuenta con una imagen del usuario.

Se replica el protocolo forense en que suelen tener una imagen indubitada y varias dudosas. De manera que es un protocolo muy complejo puesto que, en reconocimiento facial, la mayoría de sistemas funcionan mejor cuanto mayor sea el número de muestras de entrenamiento.

### 5.6.2 Resultados

En este apartado de resultados, se explican, en primer lugar, los resultados de segmentación (donde se compara el marcado manual y el automático) de la base de datos de Scface, y en segundo lugar, un análisis y una comparativa de los EERs resultantes de aplicar los extractores a cada rasgo facial (Figura 77), clasificándolos mediante un sistema basado en PCA-SVM. En la Figura 76, se muestra el usuario de la base de datos ATVS, al cual se le aplica el extractor por puntos.

Imagen original



Figura 76. Usuario de la base de datos ATVS

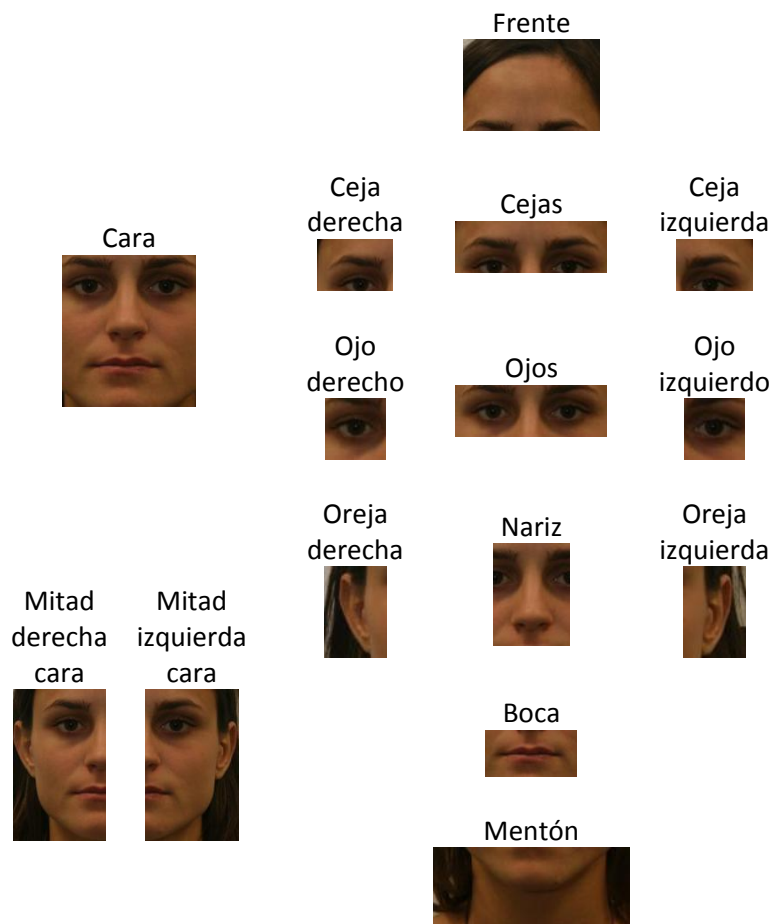
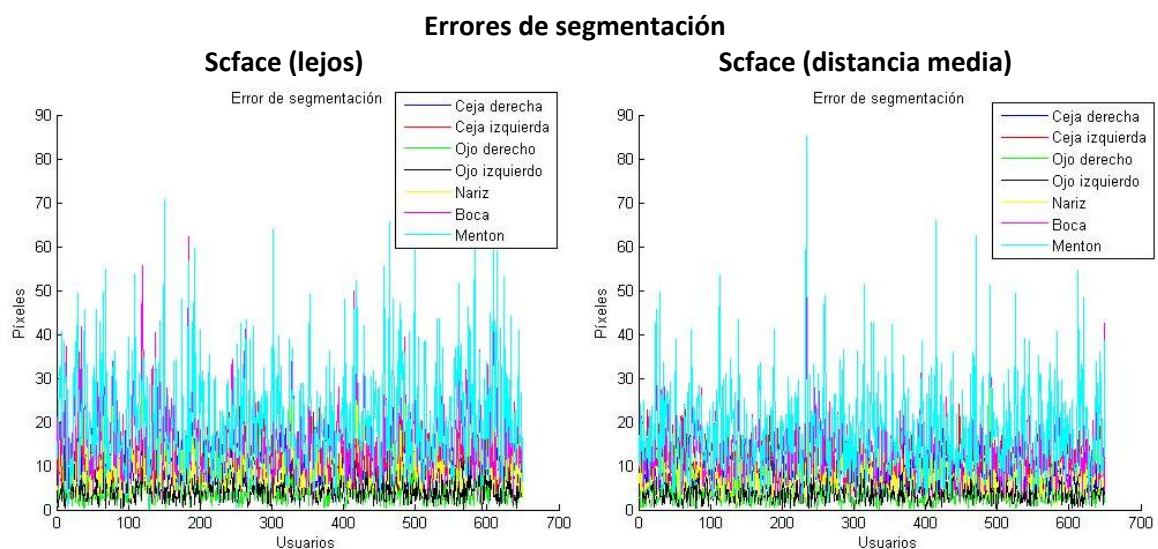


Figura 77. Rasgos faciales extraídos.

En la primera parte, como he explicado anteriormente, se van a mostrar, en la Figura 78, los errores de segmentación medios de cada rasgo (marcado manual frente al marcado automático producido por el sistema comercial), para cado uno de los usuarios que componen la base de datos Scface. Adicionalmente, se comparará dicho error dependiendo de la distancia a la que sea tomada la fotografía. Debido a que este error de segmentación es muy alto, se imposibilita realizar un reconocimiento antropométrico en este tipo de entornos, ya que los EERs son inadmisibles.



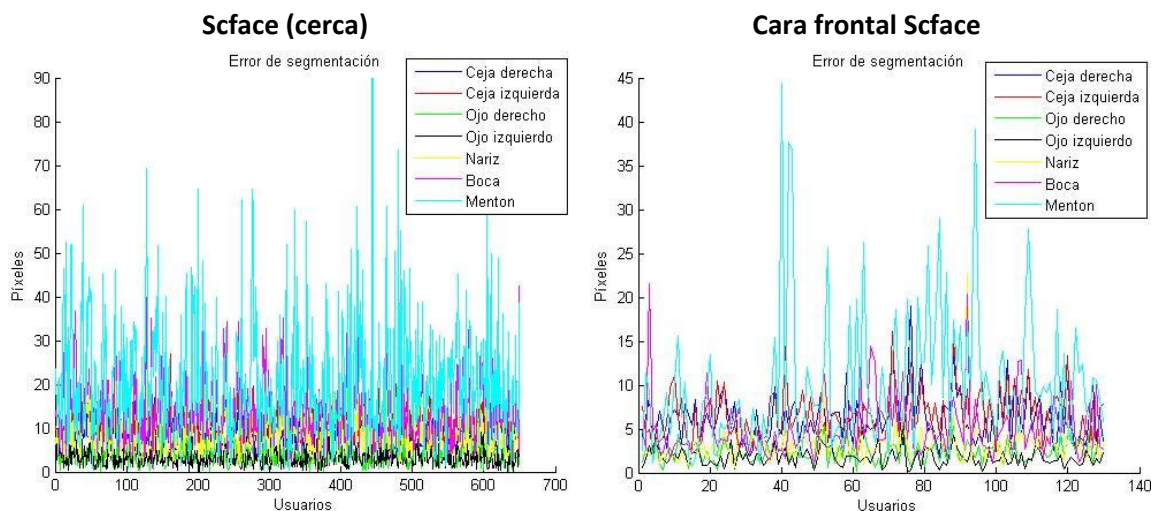


Figura 78. Error medio por rasgos y distancia.

En esta segunda parte se va a llevar a cabo el reconocimiento de patrones propiamente dicho. El reconocimiento de patrones consiste en la construcción de un sistema de clasificación a partir de un conjunto de vectores de características. En este caso, para analizar el poder discriminante de los rasgos faciales, se usará un sistema PCA-SVM.

Los vectores de características, están formados por los píxeles que componen las regiones extraídas de la cara a estudiar, es decir, la cara, las mitades izquierda y derecha de la cara, la frente, las cejas (tanto de forma conjunta como separada), los ojos (tanto de forma conjunta como separada), las orejas (de forma individual), la nariz, la boca y la barbilla.

Una vez se tienen dichos vectores de características, se hace uso de un clasificador, basado en Máquinas de Vector Soporte (SVM) [30].

Antes de la obtención de los resultados, lo primero que hay que estudiar es cuáles de las características de los precedentes vectores son las más idóneas para el reconocimiento de patrones de textura, ya que en reconocimiento de patrones ocurre lo que se conoce como “maldición de la dimensionalidad” (*Curse of Dimensionality*), la cual impone que el resultado de una clasificación no mejora necesariamente con un incremento del número de características. Para tamaños pequeños de muestra, el resultado del clasificador mejora al agregar nuevas características hasta alcanzar un máximo, para luego decaer.

Dado un número finito de ejemplos de entrenamiento, existe un número óptimo de características necesarias para clasificar. Es crucial, para ello, adoptar un esquema de selección o extracción de características [31]. En este trabajo se ha escogido el esquema **Principal component analysis, PCA**.

Por tanto, para analizar el poder discriminante de cada rasgo, se ha utilizado un sistema basado en PCA (para reducir dimensionalidades) más SVM (para clasificar los vectores de características, con kernel lineal). De forma que así se confecciona el sistema final de PCA + SVM.



El algoritmo PCA es una técnica utilizada para reducir la dimensionalidad de un conjunto de datos; esto se consigue buscando la proyección según la cual, los datos queden mejor representados en términos de mínimos cuadrados. **El número de dimensiones obtenidas con el PCA son 130, aproximadamente donde la varianza de los autovalores es del 95%.**

En la Figura 79, se representan los resultados obtenidos para ambas bases de datos, Scface y Forensic ATVS, después de haber aplicado PCA a los vectores de características (fijando 130 dimensiones) y ambos segmentadores (el basado en proporciones faciales y el basado en el punto medio de cada rasgo facial, por separado).

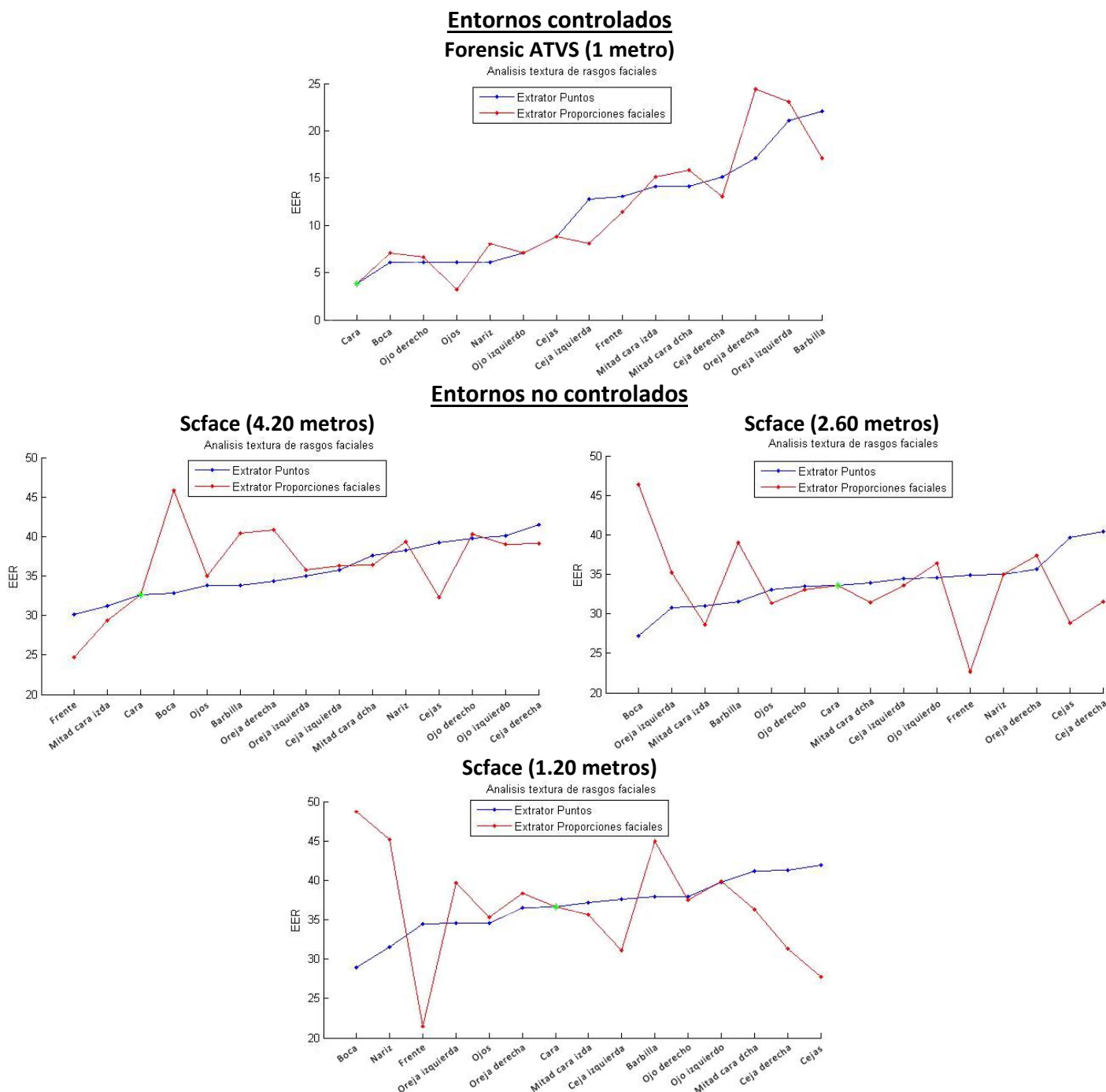


Figura 79. EER rasgos faciales mediante texturas para ambos segmentadores ordenados de mejor a peor.

En la Tabla 21, se reseñan los mejores EERs obtenidos en cada caso, clasificados por segmentador y bases de datos.

	Scface						Forensic ATVS	
Distancias	4.20 metros		2.60 metros		1.20 metros		1 metro	
Extractores	Puntos	Proporciones	Puntos	Proporciones	Puntos	Proporciones	Puntos	Proporciones
EER (%)	30.09 %	24.71 %	27.17 %	22.63 %	28.96 %	21.41 %	3.79 %	3.18 %

Tabla 21. EERs clasificados por segmentador y bases de datos.

### 5.6.3 Discusión

Una vez desarrollado el experimento que ha vertido los resultados anteriormente reflejados y, realizado el oportuno y detallado análisis de los mismos, se puede determinar en un primer lugar que el extractor basado en proporciones faciales, tiene un funcionamiento similar que el extractor basado en los puntos medios de los rasgos faciales, en el mismo tipo de entorno.

En segundo lugar y no por ello menos importante, se observa que ambos segmentadores tienen un mejor resultado en entornos controlados que en entornos no controlados, resultado no sorprendente, debido a que el marcado de puntos característicos en entornos controlados es mucho más exacto.

Otro punto a tener en cuenta es que, como se observa en la Figura 79 (entorno controlado), cuando se está en un entorno controlado y se entrena y testea el sistema con el mismo tipo de imágenes, los mejores resultados se alcanzan con la cara completa. Esto es debido a que la cara conlleva información discriminante, como la simetría entre rasgos, expresiones o gestos de la cara, cicatrices, bigote, gafas, así como otras peculiaridades que, analizando rasgo a rasgo, no se tienen en cuenta.

Finalmente, cuando se observan los resultados obtenidos en entornos no controlados, donde el entrenamiento se hace con caras frontales de buena calidad y el testeo con imágenes de webcam de baja calidad, se aprecian variaciones. En este caso, a mayor distancia, la parte de la frente adquiere un alto poder discriminativo, y según se acercan los usuarios a la cámara, la boca o la nariz obtienen los mejores resultados.



## 6. Conclusiones y trabajo futuro

---

### 6.1 Conclusiones

---

El proyecto aquí desarrollado ha tratado de describir y analizar el estudio, desarrollo, implantación y documentación de un conjunto de técnicas, basadas en documentación existente, que permiten el análisis facial tanto en entornos controlados como en entornos no controlados, utilizando para ello técnicas de estado del arte basadas en puntos característicos de la cara. Además de analizar el poder discriminante de cada rasgo.

Debido a que, actualmente, el reconocimiento facial está muy estudiado, se ha decidido orientar el proyecto a la detección y corrección de los puntos característicos de la cara, una cosa poco tratada. Como novedad, decir que estos dos importantes bloques, se aplican a los rasgos faciales de forma individual, extraídos mediante la utilización de segmentadores, en vez de a la cara como si fuese un único rasgo.

En el desarrollo de este sistema biométrico se ha propuesto un método de extracción de características basado en la detección de puntos característicos de la cara, seguido de la medida de distancias (11 en el caso de la DGGC y 82 en el caso de ATVS) entre dichos puntos. Estas distancias forman el vector de características de cada imagen.

Después se realizaron diferentes experimentos de verificación para evaluar el sistema desarrollado. Cada experimento se llevó a cabo bajo diferentes condiciones relativas a los entornos, calidad de imágenes y el número de imágenes de train y test. Cada experimento se realizó cuatro veces: i) haciendo uso de las 11 características, ii) usando sólo el conjunto de características más discriminativo (obtenido mediante el algoritmo SFFS), iii) utilizando las 82 características, y iv) usando sólo el conjunto de características más discriminativo (obtenido mediante el algoritmo SFFS); exceptuando el relativo al poder discriminante de cada rasgo, que se realizó una sola vez (con el vector de características de dimensión 130 obtenido mediante el PCA)

Conviene dejar patente, que los comparadores utilizados en el desarrollo del proyecto tienen grandes limitaciones, de hecho el SVM al tener una sola imagen de entrenamiento, se produce un sobreajuste es decir, el modelo aprendido se ajusta demasiado a los datos de entrenamiento, generalizando inadecuadamente a datos nuevos; de manera similar ocurre con el de Distancias Euclídeas con imágenes de mala calidad, donde es difícil marcar bien los puntos, igual de manera manual que de manera automática, sin duda este clasificador falla.

Los resultados obtenidos de EER en los experimentos son de interés, debido a que se han intentado buscar situaciones reales, en las cuales solo se disponen de una imagen de entrenamiento por usuario, emulando la práctica forense habitual.

Los resultados obtenidos proporcionan información sobre la influencia que tienen, sobre el rendimiento del sistema, i) las diferentes condiciones mencionadas, ii) la selección de características, y iii) el efecto de la distancia habida a la hora de la obtención de imágenes:

- El efecto de utilizar una base de datos obtenida en entornos no controlados empeora el rendimiento del sistema.
- El efecto de entrenar con imágenes de gran calidad y testear con imágenes de baja calidad también reduce el rendimiento del sistema de forma notable.
- Incorporar varios vectores de características con diferentes poses para crear los modelos de entrenamiento, es una forma de superar la falta de robustez frente a la variabilidad de pose entre los datos de entrenamiento y de test.
- El uso de un mayor número de imágenes de entrenamiento y/o test al realizar los experimentos de reconocimiento, (i) disminuye el rendimiento del sistema cuando se consideran todas las características (debido al hecho de que tener más datos de entrenamiento/test implica mayor variabilidad entre ellos en nuestros protocolos) y, (ii) lo aumenta si se realiza la selección de características (como es de esperar).
- La selección de características mejora considerablemente la el rendimiento de verificación.
- Al aplicar la selección de características, las curvas DET (gráficas que muestran el rendimiento de verificación) mejoran acercándose al origen y teniendo una pendiente más negativa, dotando al sistema de una menor (mejor) EER (Equal Error Rate, tasa de igual error en verificación) y de una mayor seguridad (la probabilidad de falsa aceptación disminuye y la de falso rechazo aumenta).

Hablando de los rasgos faciales de manera individual, se observa que cuando se realiza un análisis de textura, los EERs varían entre 3.18% y 3.79% (entornos controlados) y 21.41% y 30.09% (entornos no controlados variando la distancia entre 1.2 y 4.2 metros); mientras que cuando se realizan estudios antropométricos de éstos, los EERs varían entre 9.18% y 21.75% (entornos controlados). En entornos incontrolados, los resultados rondan el 42%.

Examinando los EERs en un entorno controlado, se observa que el mejor marcado es el manual (como era de esperar), con unos EERs que oscilan entre 8.69% (82 características) y 16.3% (11 características). El marcado automático, proporcionado por un sistema comercial llamado FaceSDK, tiene como resultados unos EERs entre el 17.1% (82 características) y 22.7% (11 características). Mejorando estos resultados, se encuentra el marcado automático desarrollado durante el proyecto, con EERs entre el 17.0% (82 características) y 22.7% (11 características).

## 6.2 Trabajo futuro

---

Como trabajo futuro, se proponen las siguientes líneas de investigación:

- La continuación del estudio antropométrico sobre bases de datos mayores, para poder analizar los resultados comprobando su potencial.
- El estudio del efecto que tiene el marcado de puntos, es decir, repercusión de los errores cometidos sobre las tasas de error del sistema desarrollado.

- El análisis detallado del poder discriminante de cada rasgo facial (cara, frente, ojos, cejas, orejas, boca, nariz y mentón) en grandes poblaciones (necesario base de datos grande)
- La adaptación del sistema aquí desarrollado para que opere, exclusivamente, con imágenes reales de baja calidad obtenidas mediante un circuito cerrado de televisión (Scface dispone de este tipo de imágenes, pero en los experimentos realizados, siempre se ha entrenado con imágenes de gran calidad).
- La fusión de algoritmos en el comparador, es decir, combinar la score obtenida mediante la realización de un estudio antropométrico y la score solución de realizar un análisis de texturas.

## 7. Referencias bibliográficas

---

- [1] [http://www.frontex.europa.eu/assets/Publications/Research/Best Practice Technical Guidelines for Automated Border Control Systems.pdf](http://www.frontex.europa.eu/assets/Publications/Research/Best_Practice_Technical_Guidelines_for_Automated_Border_Control_Systems.pdf)
- [2] [http://www.frontex.europa.eu/assets/Publications/Research/Best Practice Operational Guidelines for Automated Border Control.pdf](http://www.frontex.europa.eu/assets/Publications/Research/Best_Practice_Operational_Guidelines_for_Automated_Border_Control.pdf)
- [3] A. K. Jain, A. Ross and S. Prabhakar, "Biometrics: A Tool for Information Security", IEEE Transactions on information forensics and Security, Vol. 1, No. 2, pp. 125-143, June 2006.
- [4] Anil K. Jain, Ajay Kumar, "Biometrics of Next generation: An overview", to appear in 'Second Generation Biometrics' Springer, 2010.
- [5] R. C. Gonzalez, R. E. Woods, S. L. Eddins, "Digital Image Processing using Matlab", Prince Hall, 2005.
- [6] Gert M. Beumer, "Face recognition, a landmark tale", PhD Thesis, October 16<sup>th</sup> 2009.
- [7] Asker M. Bazer, Raymond N.J. Veldhuis, and Gerrie H. Croonen, "Likelihood Ratio-Based Detection of Facial Features", pp. 323-329, 2003.
- [8] G.M. Beumer, Q. Tao, A.M. Bazen, and R.N.J. Velhuis, "A landmark paper in face recognition", 2006.
- [9] Efraty, B.; Huang, C.; Shah, S.K.; Kakadiaris, I.A.; , "Facial landmark detection in uncontrolled conditions," Biometrics (IJCB), 2011 International Joint Conference on , vol., no., pp.1-8, 11-13 Oct. 2011.
- [10] Michal Uříčář, "Detector of facial landmark", May 13, 2011.
- [11] O. Jesorsky, K. Kirchberg, R. Frischholz, Robust Face Detection Using the Hausdorff Distance. In J. Bigun and F. Smeraldi, editors, Audio and Video based Person Authentication - AVBPA 2001, pages 90-95. Springer, 2001.
- [12] Karl Ricanek Jr and Tamirat Tesafaye, "MORPH: A Longitudinal Image Database of Normal Adult Age-Progression," IEEE 7th International Conference on Automatic Face and Gesture Recognition, Southampton, UK, April 2006, pp 341-345.
- [13] Mislav Grgic, Kresimir Delac, Sonja Grgic, SCface - surveillance cameras face database, Multimedia Tools and Applications Journal, Vol. 51, No. 3, February 2011, pp. 863-879.
- [14] Alphonse Bertillon, Alphonse Bertillon's instructions for taking descriptions for the identification of criminals, and others, by the means of anthropometric indications, 1889.
- [15] A. J. Goldstein, L. D. Harmon, and A. B. Lesk, "Identification of Human Faces", Proc. IEEE, May 1971, Vol. 59, No. 5, 748-760.
- [16] L. Sirovich and M. Kirby, "A Low-Dimensional Procedure for the Characterization of Human Faces", J. Optical Soc. Am. A, 1987, Vol. 4, No. 3, 519-524.
- [17] M. A. Turk and A. P. Pentland, "Face Recognition Using Eigenfaces", Proc. IEEE, 1991, 586-591.

- [18] Laurenz Wiskott, Jean-Marc Fellous, Norbert Krüger, Christoph Von Der Malsburg, "Face recognition by Elastic Bunch Graph Matching". IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 775-779, Julio de 1997.
- [19] T. Cootes, G. Edwards, and C. Taylor, "Active appearance models," Computer Vision-ECCV'98, p. 484, 1998.
- [20] Aditya Abhyankar and Stephanie Schuckers. Active shape models for effective iris segmentation. In SPIE 6202: Biometric Technology for Human Identification III, pages 6202:H1–H10, 2006.
- [21] Klmpak B, Grgic M, Delac K (2006) Acquisition of a face database for video surveillance research. Proceedings of the 48th international symposium ELMAR-2006 focused on multimedia signal processing and communications, Zadar, Croatia, pp 111–114.
- [22] [www.luxand.com/download/Luxand\\_FaceSDK\\_Documentation.pdf](http://www.luxand.com/download/Luxand_FaceSDK_Documentation.pdf)
- [23] ISO/IEC JTC 1/SC 37 N 504. Biometric data interchange formats part 5: Face image, in ISO, 2004.
- [24] Ye-Peng Guan, "Unsupervised human height estimation from a single image". Vol.2, No. 6, 425-430, 2009.
- [25] Yosh Jefferson, "Facial Beauty – Establishing a Universal Standar".2003.
- [26] Hatice Gunes and Massimo Piccardi, "Assessing facial beauty through proportion analysis by image processing and supervised learning", May, 2006.
- [27] Yosh Jefferson, Skeletal Types: key to unraveling the mistery of facial beauty and its significance. Published in "Journal of General Orthodontics", Volume 7, No. 2, June, 1996.
- [28] Luis Carlos Molina, Lluís Belanche, and Àngela Nebot. Feature selection algorithms: Asurvey and experimental evaluation. In Proceedings of the 2002 IEEE International Conference on Data Mining, ICDM '02, pages 306–, Washington, DC, USA, 2002. IEEE Computer Society.
- [29] Pedro Tomé, Luis Blázquez, Rubén Vera-Rodriguez, Julián Fierrez, Javier Ortega-García, Nicomedes Expósito y Patricio Lestón. "Understanding the discrimination power of facial regions in forensic casework". **International Workshop on Biometrics and Forensics** (IWBF), Abril 4-5, 2013, Lisboa, Portugal.
- [30] Corinna Cortes and Vladimir Vapnik. "Support-Vector Networks". Machine Learning. Vol. 20, Num. 3, pp. 273-297, September 1995.
- [31] P. Pudil, J. Novovicová, and J. Kittler. Floating Search Methods in Feature Selection. Pattern Recognition Letters, 15(11):1119 – 1125, 1994.

# A. Presupuesto

---

<b>1) Ejecución material</b>	
• Compra de ordenador personal (Software incluido)	2000 €
• Alquiler de impresora láser durante 6 meses	50 €
• Material de oficina	150 €
• Total de ejecución material	2200 €
<b>2) Gastos generales</b>	
• 16 % sobre Ejecución Material	352 €
<b>3) Beneficio industrial</b>	
• 6 % sobre Ejecución Material	132 €
<b>4) Honorarios Proyecto</b>	
• 1500 horas a 15 € / hora	22500 €
<b>5) Material fungible</b>	
• Gastos de impresión	60 €
• Encuadernación	200 €
• Total material fungible	260 €
<b>6) Subtotal del presupuesto</b>	
• Subtotal presupuesto	24960 €
<b>7) I.V.A. aplicable</b>	
• 21 % Subtotal Presupuesto	3993,6 €
<b>8) Total presupuesto</b>	
• Total presupuesto	28953,6 €

Madrid, Enero 2013

El Ingeniero Jefe de Proyecto

Fdo.: Luis Blázquez Pérez  
Ingeniero Superior de Telecomunicación

## B. Pliego de condiciones

---

Este documento contiene las condiciones legales que guiarán la realización, en este proyecto, de un **Reconocimiento Facial Basado en Puntos Característicos de la Cara en entornos no controlados**. En lo que sigue, se supondrá que el proyecto ha sido encargado por una empresa cliente a una empresa consultora con la finalidad de realizar dicho sistema. Dicha empresa ha debido desarrollar una línea de investigación con objeto de elaborar el proyecto. Esta línea de investigación, junto con el posterior desarrollo de los programas está amparada por las condiciones particulares del siguiente pliego.

Supuesto que la utilización industrial de los métodos recogidos en el presente proyecto ha sido decidida por parte de la empresa cliente o de otras, la obra a realizar se regulará por las siguientes:

### Condiciones generales

1. La modalidad de contratación será el concurso. La adjudicación se hará, por tanto, a la proposición más favorable sin atender exclusivamente al valor económico, dependiendo de las mayores garantías ofrecidas. La empresa que somete el proyecto a concurso se reserva el derecho a declararlo desierto.

2. El montaje y mecanización completa de los equipos que intervengan será realizado totalmente por la empresa licitadora.

3. En la oferta, se hará constar el precio total por el que se compromete a realizar la obra y el tanto por ciento de baja que supone este precio en relación con un importe límite si este se hubiera fijado.

4. La obra se realizará bajo la dirección técnica de un Ingeniero Superior de Telecomunicación, auxiliado por el número de Ingenieros Técnicos y Programadores que se estime preciso para el desarrollo de la misma.

5. Aparte del Ingeniero Director, el contratista tendrá derecho a contratar al resto del personal, pudiendo ceder esta prerrogativa a favor del Ingeniero Director, quien no estará obligado a aceptarla.

6. El contratista tiene derecho a sacar copias a su costa de los planos, pliego de condiciones y presupuestos. El Ingeniero autor del proyecto autorizará con su firma las copias solicitadas por el contratista después de confrontarlas.

7. Se abonará al contratista la obra que realmente ejecute con sujeción al proyecto que sirvió de base para la contratación, a las modificaciones autorizadas por la superioridad o a las órdenes que con arreglo a sus facultades le hayan comunicado por escrito al Ingeniero Director de obras siempre que dicha obra se haya ajustado a los preceptos de los pliegos de condiciones, con arreglo a los cuales, se harán las modificaciones y la valoración de las diversas unidades sin que el importe total pueda exceder de los presupuestos aprobados. Por consiguiente, el número de unidades que se consignan en el proyecto o en el presupuesto, no podrá servirle de fundamento para entablar reclamaciones de ninguna clase, salvo en los casos de rescisión.



8. Tanto en las certificaciones de obras como en la liquidación final, se abonarán los trabajos realizados por el contratista a los precios de ejecución material que figuran en el presupuesto para cada unidad de la obra.

9. Si excepcionalmente se hubiera ejecutado algún trabajo que no se ajustase a las condiciones de la contrata pero que sin embargo es admisible a juicio del Ingeniero Director de obras, se dará conocimiento a la Dirección, proponiendo a la vez la rebaja de precios que el Ingeniero estime justa y si la Dirección resolviera aceptar la obra, quedará el contratista obligado a conformarse con la rebaja acordada.

10. Cuando se juzgue necesario emplear materiales o ejecutar obras que no figuren en el presupuesto de la contrata, se evaluará su importe a los precios asignados a otras obras o materiales análogos si los hubiere y cuando no, se discutirán entre el Ingeniero Director y el contratista, sometiéndolos a la aprobación de la Dirección. Los nuevos precios convenidos por uno u otro procedimiento, se sujetarán siempre al establecido en el punto anterior.

11. Cuando el contratista, con autorización del Ingeniero Director de obras, emplee materiales de calidad más elevada o de mayores dimensiones de lo estipulado en el proyecto, o sustituya una clase de fabricación por otra que tenga asignado mayor precio o ejecute con mayores dimensiones cualquier otra parte de las obras, o en general, introduzca en ellas cualquier modificación que sea beneficiosa a juicio del Ingeniero Director de obras, no tendrá derecho sin embargo, sino a lo que le correspondería si hubiera realizado la obra con estricta sujeción a lo proyectado y contratado.

12. Las cantidades calculadas para obras accesorias, aunque figuren por partida alzada en el presupuesto final (general), no serán abonadas sino a los precios de la contrata, según las condiciones de la misma y los proyectos particulares que para ellas se formen, o en su defecto, por lo que resulte de su medición final.

13. El contratista queda obligado a abonar al Ingeniero autor del proyecto y director de obras así como a los Ingenieros Técnicos, el importe de sus respectivos honorarios facultativos por formación del proyecto, dirección técnica y administración en su caso, con arreglo a las tarifas y honorarios vigentes.

14. Concluida la ejecución de la obra, será reconocida por el Ingeniero Director que a tal efecto designe la empresa.

15. La garantía definitiva será del 4% del presupuesto y la provisional del 2%.

16. La forma de pago será por certificaciones mensuales de la obra ejecutada, de acuerdo con los precios del presupuesto, deducida la baja si la hubiera.

17. La fecha de comienzo de las obras será a partir de los 15 días naturales del replanteo oficial de las mismas y la definitiva, al año de haber ejecutado la provisional, procediéndose si no existe reclamación alguna, a la reclamación de la fianza.

18. Si el contratista al efectuar el replanteo, observase algún error en el proyecto, deberá comunicarlo en el plazo de quince días al Ingeniero Director de obras, pues transcurrido ese plazo será responsable de la exactitud del proyecto.

19. El contratista está obligado a designar una persona responsable que se entenderá con el Ingeniero Director de obras, o con el delegado que éste designe, para todo relacionado con ella. Al ser el Ingeniero Director de obras el que interpreta el proyecto, el contratista deberá consultarle cualquier duda que surja en su realización.

20. Durante la realización de la obra, se girarán visitas de inspección por personal facultativo de la empresa cliente, para hacer las comprobaciones que se crean oportunas. Es obligación del contratista, la conservación de la obra ya ejecutada hasta la recepción de la misma, por lo que el deterioro parcial o total de ella, aunque sea por agentes atmosféricos u otras causas, deberá ser reparado o reconstruido por su cuenta.

21. El contratista, deberá realizar la obra en el plazo mencionado a partir de la fecha del contrato, incurriendo en multa, por retraso de la ejecución siempre que éste no sea debido a causas de fuerza mayor. A la terminación de la obra, se hará una recepción provisional previo reconocimiento y examen por la dirección técnica, el depositario de efectos, el interventor y el jefe de servicio o un representante, estampando su conformidad el contratista.

22. Hecha la recepción provisional, se certificará al contratista el resto de la obra, reservándose la administración el importe de los gastos de conservación de la misma hasta su recepción definitiva y la fianza durante el tiempo señalado como plazo de garantía. La recepción definitiva se hará en las mismas condiciones que la provisional, extendiéndose el acta correspondiente. El Director Técnico propondrá a la Junta Económica la devolución de la fianza al contratista de acuerdo con las condiciones económicas legales establecidas.

23. Las tarifas para la determinación de honorarios, reguladas por orden de la Presidencia del Gobierno el 19 de Octubre de 1961, se aplicarán sobre el denominado en la actualidad "Presupuesto de Ejecución de Contrata" y anteriormente llamado "Presupuesto de Ejecución Material" que hoy designa otro concepto.

### **Condiciones particulares**

La empresa consultora, que ha desarrollado el presente proyecto, lo entregará a la empresa cliente bajo las condiciones generales ya formuladas, debiendo añadirse las siguientes condiciones particulares:

1. La propiedad intelectual de los procesos descritos y analizados en el presente trabajo, pertenece por entero a la empresa consultora representada por el Ingeniero Director del Proyecto.

2. La empresa consultora se reserva el derecho a la utilización total o parcial de los resultados de la investigación realizada para desarrollar el siguiente proyecto, bien para su publicación o bien para su uso en trabajos o proyectos posteriores, para la misma empresa cliente o para otra.

3. Cualquier tipo de reproducción aparte de las reseñadas en las condiciones generales, bien sea para uso particular de la empresa cliente, o para cualquier otra aplicación, contará con autorización expresa y por escrito del Ingeniero Director del Proyecto, que actuará en representación de la empresa consultora.

4. En la autorización se ha de hacer constar la aplicación a que se destinan sus reproducciones así como su cantidad.

5. En todas las reproducciones se indicará su procedencia, explicitando el nombre del proyecto, nombre del Ingeniero Director y de la empresa consultora.

6. Si el proyecto pasa la etapa de desarrollo, cualquier modificación que se realice sobre él, deberá ser notificada al Ingeniero Director del Proyecto y a criterio de éste, la empresa consultora decidirá aceptar o no la modificación propuesta.

7. Si la modificación se acepta, la empresa consultora se hará responsable al mismo nivel que el proyecto inicial del que resulta el añadirla.

8. Si la modificación no es aceptada, por el contrario, la empresa consultora declinará toda responsabilidad que se derive de la aplicación o influencia de la misma.

9. Si la empresa cliente decide desarrollar industrialmente uno o varios productos en los que resulte parcial o totalmente aplicable el estudio de este proyecto, deberá comunicarlo a la empresa consultora.

10. La empresa consultora no se responsabiliza de los efectos laterales que se puedan producir en el momento en que se utilice la herramienta objeto del presente proyecto para la realización de otras aplicaciones.

11. La empresa consultora tendrá prioridad respecto a otras en la elaboración de los proyectos auxiliares que fuese necesario desarrollar para dicha aplicación industrial, siempre que no haga explícita renuncia a este hecho. En este caso, deberá autorizar expresamente los proyectos presentados por otros.

12. El Ingeniero Director del presente proyecto, será el responsable de la dirección de la aplicación industrial siempre que la empresa consultora lo estime oportuno. En caso contrario, la persona designada deberá contar con la autorización del mismo, quien delegará en él las responsabilidades que ostente.

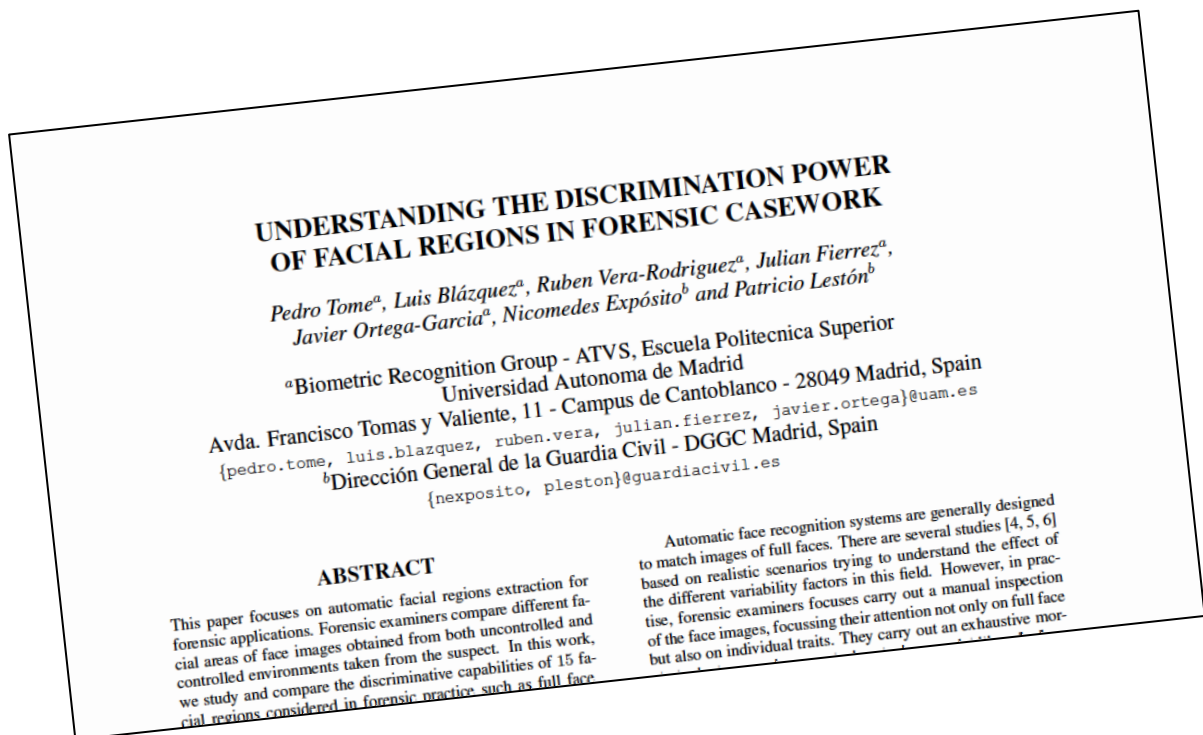
## C. Publicaciones

Durante la realización del proyecto, se ha llevado a cabo una publicación, aceptada en el **International Workshop on Biometrics and Forensics (IWBF)**, el cual se celebrará en Lisboa (Abril del 2013).

**Title:** Understanding the Discrimination Power of Facial Regions in Forensic Casework.

**Authors:** Pedro Tomé, Luis Blázquez, Rubén Vera-Rodriguez, Julián Fierrez, Javier Ortega-García, Nicomedes Expósito y Patricio Lestón.

**Published in:** **International Workshop on Biometrics and Forensics (IWBF)**, Abril 4-5, 2013, Lisboa, Portugal.



## UNDERSTANDING THE DISCRIMINATION POWER OF FACIAL REGIONS IN FORENSIC CASEWORK

Pedro Tome<sup>a</sup>, Luis Blázquez<sup>a</sup>, Ruben Vera-Rodriguez<sup>a</sup>, Julian Fierrez<sup>a</sup>,  
Javier Ortega-Garcia<sup>a</sup>, Nicomedes Expósito<sup>b</sup> and Patricio Lestón<sup>b</sup>

<sup>a</sup>Biometric Recognition Group - ATVS, Escuela Politécnica Superior  
Universidad Autónoma de Madrid

Avda. Francisco Tomas y Valiente, 11 - Campus de Cantoblanco - 28049 Madrid, Spain

{pedro.tome, luis.blazquez, ruben.vera, julian.fierrez, javier.ortega}@uam.es

<sup>b</sup>Dirección General de la Guardia Civil - DGGC Madrid, Spain

{nexpósito, pleston}@guardiacivil.es

### ABSTRACT

This paper focuses on automatic facial regions extraction for forensic applications. Forensic examiners compare different facial areas of face images obtained from both uncontrolled and controlled environments taken from the suspect. In this work, we study and compare the discriminative capabilities of 15 facial regions considered in forensic practice such as full face, nose, eye, eyebrow, mouth, etc. This study is useful because it can statistically support the current practice of forensic facial comparison. It is also of interest to biometrics because a more robust general-purpose face recognition system can be built by fusing the similarity scores obtained from the comparison of different individual parts of the face. To analyse the discrimination power of each facial region, we have randomly defined three population subsets of 200 European subjects (male, female and mixed) from MORPH database. First facial landmarks are automatically located, checked and corrected and then 15 forensic facial regions are extracted and considered for the study. In all cases, the performance of the full face (faceISOV region) is higher than the one achieved for the rest of facial regions. It is very interesting to note that the nose region has a very significant discrimination efficiency by itself and similar to the full face performance.

**Index Terms**— Forensic, biometrics, face recognition, facial regions, forensic casework.

### 1. INTRODUCTION

Automatic face recognition has been extensively researched over the past two decades. This growth is due to its easy acquisition and its important role in a growing number of application domains, including access control, video surveillance, and its wide use in government issued identity documents (e.g., passport and driving's license) [1].

An area where these kinds of systems have obtained an increased emphasis is the forensic field [2]. Forensic science analyses data collected by law enforcement agencies in order to prove or disprove the guilt of a suspect with high confidence under the legal system.

While DNA and fingerprint forensic identification are two of the most reliable and available identification methods in forensic science, automatic face recognition technology needs to improve the set of available tools to determine a person's

identity, particularly from video surveillance imagery. Such progress for forensic face recognition is one of the goals of the FBI's Next Generation Identification program [3].

Automatic face recognition systems are generally designed to match images of full faces. There are several studies [4, 5, 6, 7] based on realistic scenarios trying to understand the effect of the different variability factors in this field. However, in practice, forensic examiners focus carry out a manual inspection of the face images, focussing their attention not only on full face but also on individual traits. They carry out an exhaustive morphological comparison, analysing the intra-variability of a face, trait by trait on nose, mouth, eyebrows, etc., even examining soft traits such as marks, moles, wrinkles, etc.

As Anil K. Jain describes as future work in [8, 9, 2], a component-based system (facial regions-based system) where given a probe face (partial or complete) and the matching and retrieval would be performed for each facial component (eyes, nose, mouth, chin, eyebrows, etc.) would be of great value to forensic investigators.

There are some previous works where this effect is studied [10, 11, 12, 13, 14, 15] but these papers focus their attention just in some common traits. In this work, we have extracted facial components (called from now facial regions) following forensic protocols from law enforcement agencies, allowing us to study the discriminative power of different facial regions individually. In particular we address in this paper the problem of finding the most discriminative areas of the face for recognition. To the best of our knowledge, this has not been done by any of the existing face recognition approaches directly.

Understanding the discrimination power of different facial regions on a wide population have some remarkable benefits such as allowing investigators to specify a particular region of the face prevents incomplete, noisy, and missing regions from degrading the recognition accuracy. Further, a better understanding of facial regions-based face recognition should facilitate the study of individuality models.

The main contribution of the paper is an experimental study of the discriminative power of different forensic facial regions on a wide population. Additionally, we propose a novel framework for facial regions extraction useful for controlled and uncontrolled scenarios.

The rest of the paper is organized as follows. In Section 2, we provide an overview of the automatic facial region extraction procedure. Section 3 presents the analysis of the extracted

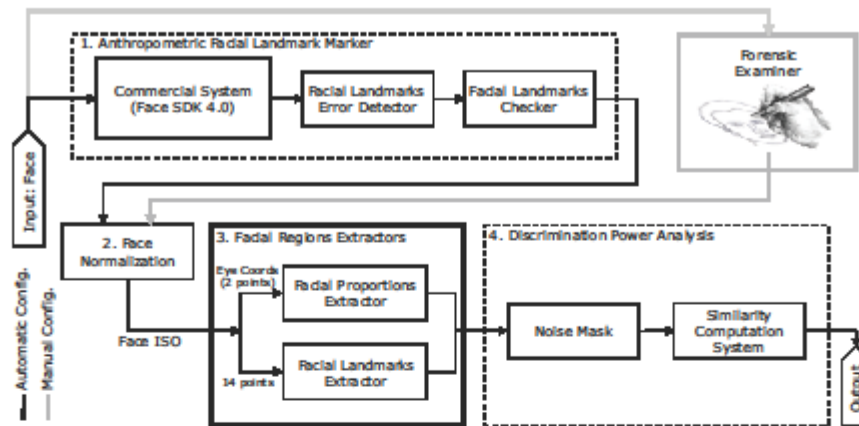


Figure 1: Experimental framework diagram description.

different facial regions defining the database used, the experimental protocol followed, the feature extraction and classification used and the experimental results achieved. We conclude in Section 4 with a discussion and summary of our work.

## 2. FACIAL REGIONS EXTRACTION

This section describes the experimental framework developed to extract the different forensic facial regions analysed in this work.

The traditional inspection procedure of the law enforcement agencies carried out by forensic examiners is mainly based on manual and individual skills of the human examiners using some general image processing tools. Automatic approaches of image processing could help the examiners to reduce the human subjective decisions, reaching higher precisions. In this sense, we have developed a useful tool able to extract different facial regions as summarized in Fig. 1.

The presented experimental framework has two different configurations: automatic and manual in order to find the facial landmarks for extraction of the facial regions. Automatic configuration uses a commercial system<sup>1</sup> that provide 65 facial points of which only 13 are used. These 13 facial landmark plus a new point that indicate the top of the head (defined by us) (see Fig. 2) are used as inputs to a facial landmarks error detector developed by us based on distances, angles and symmetries between these points. This system allows us to know which facial landmark is correctly located and which of them needs to be corrected. On the other hand, the location of these facial landmarks could be done manually by a forensic examiner.

After a correct facial landmark location, faces are normalized based on the ISO standard [16] with an inter-pupillary pixel distance (IPD) of 75 pixels. Therefore, facial regions can be extracted with a standard size for all faces.

In our approach we have implemented two different facial region extractors: *i)* based on human facial proportions and *ii)* based on facial landmarks. The first one extracts the facial area of interest of the face (eyebrows, eyes, nose, mouth, etc.) us-

ing just the two eyes coordinates, following single facial proportions rules with regarding to IPD distance. The mentioned extractor would be of interest in challenges uncontrolled scenarios where landmarks are very difficult to be extracted automatically. On the other hand, the second extractor, based on facial landmarks correctly located, allows to extract the facial regions with high precision.

The experimental framework implemented allows the extraction of 15 different facial regions as can be seen in Fig. 2. The election of these 15 regions is based on protocols from Spanish Guardia Civil [17] and NFI [18] forensic face recognition protocols.

## 3. FACIAL REGIONS ANALYSIS

This section describes how facial regions extracted from a face are analysed in order to evaluate their discriminative power. Firstly, the database and the experimental protocol adopted for this work are presented. Then, the feature extraction and classification will be described and finally, the experimental results will be detailed.

### 3.1. Database

The experiments are carried out on a subset of MORPH Non-Commercial Release [19] database. MORPH data corpus contains 55.000 frontal face images from more than 13.000 subjects, acquired from 2003 to late 2007. The distribution of ages ranges from 16 to 77 with an average age of 33. The average number of images per individual is 4 and the average time between pictures is 164 days, with the minimum being 1 day and the maximum being 1.681 days. The MORPH database is divided in 5 subsets named: *i)* african, *ii)* european, *iii)* asian, *iv)* hispanic and *v)* other.

The subset "european" comprised of 2.704 subjects (2.070 males plus 634 females) has been selected for these experiments. Fig. 2 shows an example.

<sup>1</sup>Luxand, Inc. <http://www.luxand.com>.



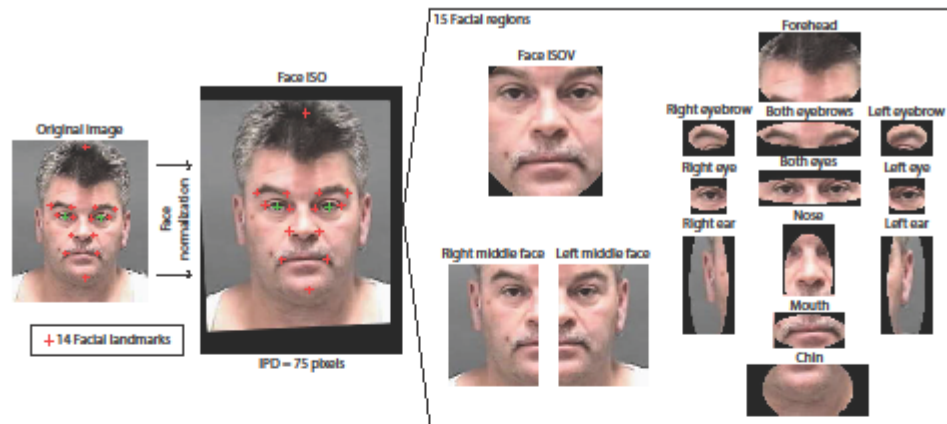


Figure 2: Summary diagram of forensic facial regions extraction.

### 3.2. Experimental Protocol

For the experimental work of this paper we discarded those subjects with less than three images and chose three images per subject with the smallest gap between acquisitions in order to reduce the lapse effect.

Then using this selection, three population sets were randomly chosen in order to analyse the discrimination power of each facial region on different populations: i) 200female, ii) 200male and iii) 200mix (100 male+100 female).

The performance of the evaluated similarity system is computed using the experimental protocol described. Each population subset of 200 subjects with 3 face images is divided into: i) a training set comprising the first sample (enrolment template); and ii) an evaluation set comprised of the rest two images available for each subject.

### 3.3. Feature Extraction and Classification

Regarding feature extraction and classification, a system based on PCA-SVM was adopted to compute the discrimination power between different facial regions. Different noise masks were applied to each facial region (Fig 2) (e.g.  $75 \times 101$  (width  $\times$  height) for nose region). Then, PCA was applied to each facial region over the training set retainin 96% of variance. This leads to a system where the original image space (e.g. of 7.575 dimensions for nose region) is reduced to 200 dimensions. Similarity scores are computed in this PCA vector space using a SVM classifier with linear kernel.

### 3.4. Experimental Results

This section describes the experimental analysis of individual features of each facial region and their discrimination power (represented by EER) over the different 3 population datasets. Results are shown using ROC curves with EERs (in %).

The discrimination power of each defined forensic facial region for the three studied population datasets is presented in Fig. 3. As can be seen, doing a global analysis, faceISOV region reaches the highest performance compared to the other facial regions, followed by the nose and middle faces regions. It is

worth highlighting that the faceISOV and middle faces include other facial regions considered. However, the nose region does not, hence it is important to remark that the nose region has a very high and important discrimination power with respect to the other regions of the face. Ranking the remaining facial parts regarding their discrimination power, the eye regions come next, followed by eyebrows, mouth or chin. The worst results were obtained for the chin, which could be explained due to difficulty to locate the corresponding landmark. As it was expected, ears achieved worse results due to the common occlusion by hair and the pose. It is important to note that mouth region achieves poor performance, it could be due to variability having a not neutral pose: open, closed, smiling, etc.

As can be seen in Fig. 3 (middle and bottom), faceISOV for male and female populations has more or less the same performance, but in general discriminative results for the male population were better than female, due to less variability.

## 4. CONCLUSIONS

In the present work, an experimental framework of extraction of different facial regions of a face has been presented and used to understand their discrimination power. The discrimination efficiency of each facial region has been studied considering three different populations provided by a subset of MORPH database. In all cases, the performance of the full face named faceISOV region is higher than the one achieved in the rest of facial regions. But in a real forensic scenario, partial faces are considered very often to the study of the discrimination power of the different regions due to occlusions or other factors, hence this individualized study is very useful for forensic face recognition. Furthermore, the nose region has a very significant discrimination efficiency by itself and similar to full face performance. There are notable differences between male and female performances on different facial regions and in general men achieve better discriminative results for their facial regions compared to women, most likely due to less variability of appearance. This work highlights the benefits of adequate analysis of facial regions from a face in order to better understand the facial intra-variability.



## 5. ACKNOWLEDGEMENT

P. Tome is supported by a FPU Fellowship from Univ. Autónoma de Madrid. This work has been partially supported by contract with Spanish Guardia Civil and projects BBfor2 (FP7-ITN-238803), Bio-Challenge (TEC2009-11186), Contexts (S2009/TIC-1485), TeraSense (CSD2008-00068) and "Cátedra UAM-Telefónica".

## 6. REFERENCES

- [1] Anil K. Jain, Arun A. Ross, and Karthik Nandakumar, *Introduction to Biometrics*, Springer Science+Business Media, LLC, 2011.
- [2] Anil K. Jain, Brendan Klare, and Unsang Park, "Face recognition: Some challenges in forensics," in *FG*, 2011, pp. 726–733, IEEE.
- [3] Next Generation Identification, "http://www.fbi.gov/about-us/cj/is/tinge/sprints/biometrics/ngid2,".
- [4] Xiaoheng Zhang and Yongsheng Gao, "Face recognition across pose: A review," *Pattern Recognition*, pp. 2876–2896, 2009.
- [5] Pedro Tome, Julian Fierrez, M. C. Fairhurst, and Javier Ortega-Garcia, "Acquisition scenario analysis for face recognition at a distance," in *6th International Symposium on Visual Computing (ISVC)*, November–December 2010.
- [6] Stan Z. Li and Anil K. Jain, Eds., *Handbook of Face Recognition*, 2nd Edition, Springer, 2011.
- [7] Pedro Tome, Ruben Vera-Rodriguez, Julian Fierrez, and Javier Ortega-Garcia, "Variability compensation using nap for unconstrained face recognition," in *Proc. 10th International Conference on Practical Applications of Agents and Multi-Agent Systems (PRAAS'12)*, March 2012, vol. 151, pp. 129–139, Springer.
- [8] D. Meuwly and R. Veldhuis, "Forensic biometrics: From two communities to one discipline," in *Biometrics Special Interest Group (BIOSIG), 2012 BIOSIG - Proceedings of the International Conference of the*, sept. 2012, pp. 1–12.
- [9] Anil K. Jain, Brendan Klare, and Unsang Park, "Face matching and retrieval in forensics applications," *IEEE Multimed.*, vol. 19, no. 1, pp. 20, 2012.
- [10] Massimo Tistarelli, "Active/space-variant object recognition," *Image and Vision Computing*, vol. 13, no. 3, pp. 215–226, 1995.
- [11] Bernd Heisele, Thomas Serre, and T. Poggio, "A component-based framework for face detection and identification," *Int. J. Computer Vision*, vol. 74, no. 2, pp. 167–181, Aug. 2007.
- [12] Fayin Li, Harry Wechsler, and Massimo Tistarelli, "Robust fusion using boosting and transduction for component-based face recognition," in *10th International Conference on Computer, Automation, Robotics and Vision (ICARCV 2008, Hanoi, Vietnam, 17-20 December 2008, Proceedings 2008)*, pp. 434–439, IEEE.
- [13] Shalini Gupta, Mia K. Markey, and Alan C. Bovik, "Anthropometric 3d face recognition," *Int. J. Comput. Vision*, vol. 90, no. 3, pp. 331–349, Dec. 2010.
- [14] Tauseef Ali, Pedro Tome, Julian Fierrez, Ruben Vera-Rodriguez, Luk Spenswens, and Raymond Veldhuis, "A study of identification performance of facial regions from cctv images," in *5th International Workshop on Computational Forensics, IWCF 2012*, November 2012.
- [15] Omar Ocegueda, Shishir K. Shah, and Ioannis A. Kakadiaris, "Which parts of the face give out your identity?," in *CVPR*, 2011, pp. 641–648, IEEE.
- [16] ISO/IEC JTC 1/SC 37 N 504, "Biometric data interchange formats part 5: Face image," in *International Organization for Standardization*, 2004.
- [17] Spanish Guardia Civil (DGGC), "http://www.guardiacivil.es/,".
- [18] Netherlands Forensic Institute (NFI), "http://www.forensicinstitute.nl/,".
- [19] K. Ricanek and T. Tesafaye, "Morph: a longitudinal image database of normal adult age-progression," in *Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on*, april 2006, pp. 341–345.

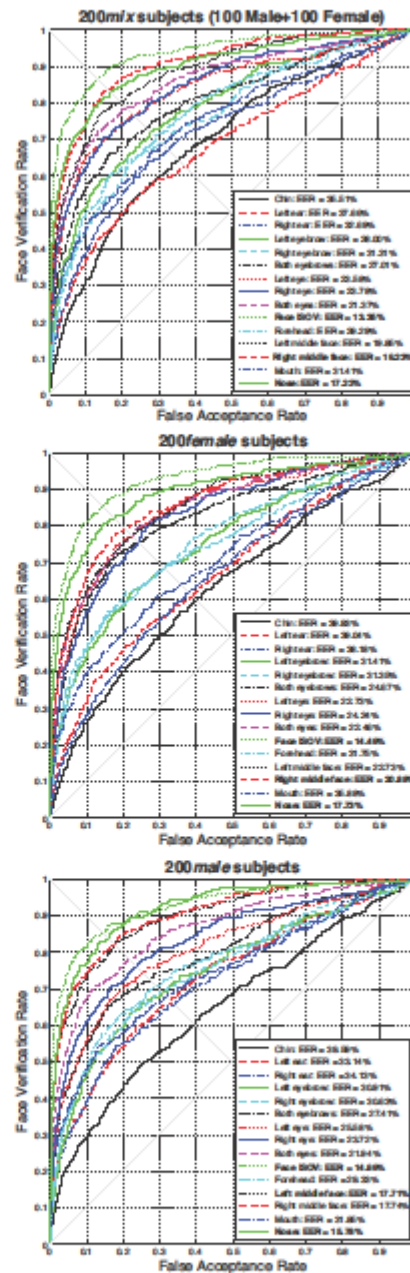


Figure 3: ROC curves showing verification performance of different facial regions (highlighting the best three regions) obtained for the three population sets: (top) 200mix, (middle) 200female and (bottom) 200male.